STAT 486: Market Models

April 6, 2018

Mini Project 4: Max-Median Rule

Group 7 Members:

Hongyu Mao

Mengyuan Li

Ningshan (Sasha) Cui

Qingyi (Jessica) Chen

Yuanrui (Rylie) Gao

CONTENTS

INTRODUCTION

In this project, we investigate the effectiveness of the Max-Median Rule, a methodology for portfolio management, on S&P 500 companies from 1970 to the latest year. We use S&P 500 as a benchmark for comparison. In addition, we improve the portfolio selection methods in three ways, keeping the same portfolio size. We utilize the mean selection method, mode selection method and kernel selection method, which is inspired in a Time Series class. The purposes for improvement are to maximize the portfolio annual return and to decrease the variance between the maximum return and minimum return. We use R to select proper stocks and perform analysis. Comparing these three methods, we conclude that Max-Median Rule is a little riskier comparing to other selection methods.

S&P500

Since the return is comparing to S&P 500. Knowing the components of S&P 500 will help us better understand the results. The Standard & Poor's 500 is an American stock market index based on the market capitalizations of 500 large companies, having common stock listed on the NYSE or NASDAQ. There are primarily eight criterion for choosing a stock: market capitalization, liquidity, domicile, public float, sector classification, financial viability, and length of time publicly traded and stock exchange.

Data Collection Process

For the listed S&P 500 companies on Canvas, we collect the daily stock returns from CRSP Daily Stock Return. The year range is from 1979 to 2016.

Max-Median Rule

Methodology of Max-Median Rule

1. Collect the first year daily price for all the stocks in S&P 500.
2. For each stock, calculate the day-to-day ratios (ratio = price today / price in the previous day) across one year, then get the median ratio for this year. Get 500 medians for 500 stocks.
3. Choose 20 stocks with highest median ratios, invest on these stocks.
4. Choose the stocks with 20 highest median of daily returns, the invest on the stocks in median/mean weighted way.

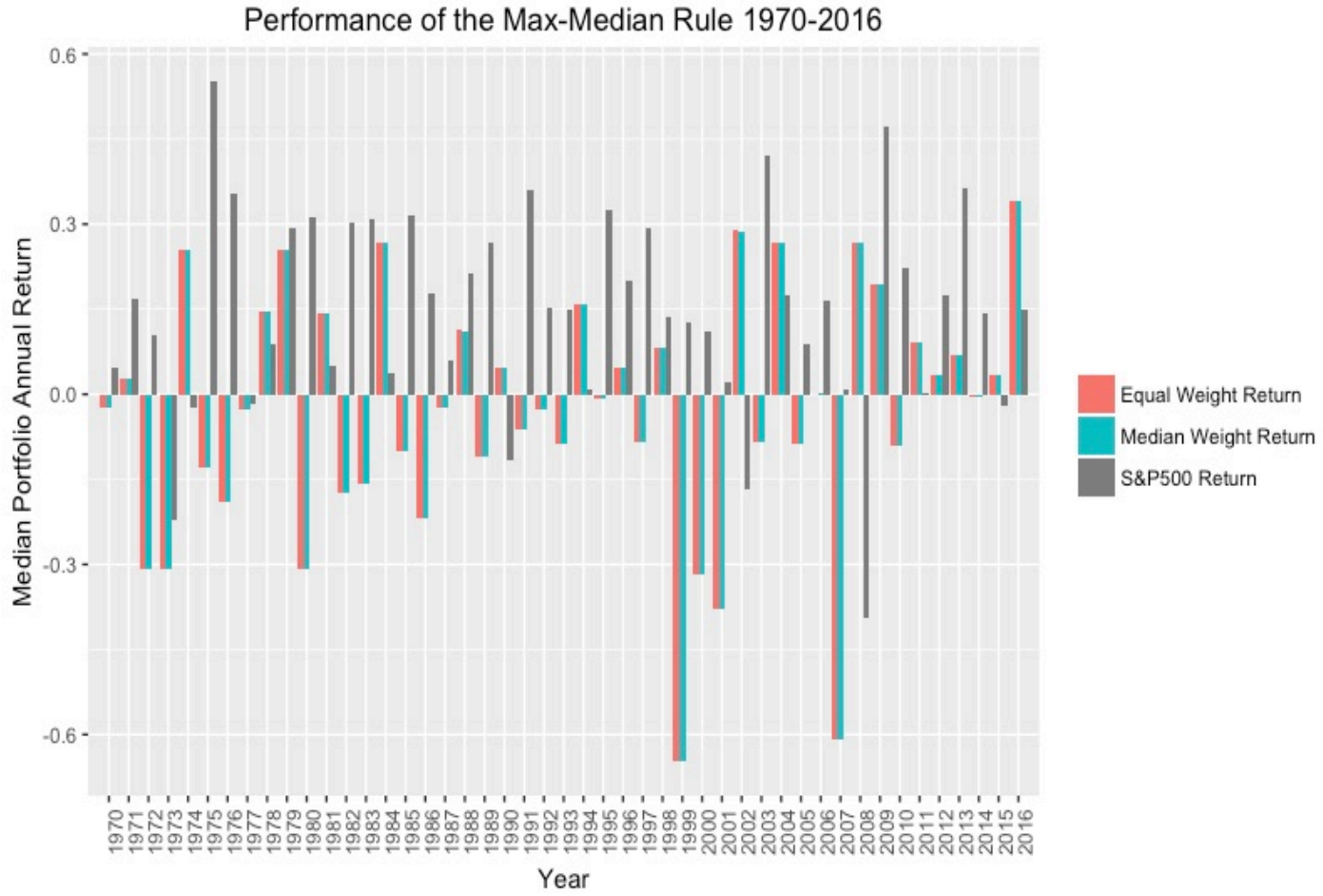

Figure 1

On this graph, the portfolio annual returns vibrate from -0.6 to +0.5. For most of the time, it is above 0, which means positive annual returns. However, for Year 1997 and Year 2007, the returns are both extremely negative. This might be due to the financial crisis from 1997-1999 and from 2007-2008. Even though the negative returns are associated with the entire financial market, we can still make conclusion that the Max-Median Rule cannot always guarantee positive and outstanding portfolio returns.

Note: To calculate the return of the 20-stock portfolio created by Max-Median rule every year, we use the following formula:

$$\frac{\sum \text{every stock's price on the last trading day} - \sum \text{every stock's price on the first trading day}}{\sum \text{every stock's price on the first trading day}}$$

However, we found out that under this formula, our Max-Median result could not outperform S&P 500. Then, we tried different formulas such as CAGR formula, but it didn't work, neither. Even after consulting our Teaching Assistant, we could not figure out what factor made out result underperform the market.

Table 1 shows the returns for the Max-Median portfolio from 1970-2016.

Year	Equal-weighted	Median-weighted	Year	Equal-weighted	Median-weighted
1970	-0.02	-0.02	**1994**	0.16	0.16
1971	0.03	0.03	**1995**	-0.01	-0.01
1972	-0.31	-0.31	**1996**	0.05	0.05
1973	-0.31	-0.31	**1997**	-0.08	-0.08
1974	0.25	0.25	**1998**	0.08	0.08
1975	-0.13	-0.13	**1999**	-0.65	-0.65
1976	-0.19	-0.19	**2000**	-0.32	-0.32
1977	-0.03	-0.03	**2001**	-0.38	-0.38
1978	0.15	0.15	**2002**	0.29	0.29
1979	0.26	0.26	**2003**	-0.08	-0.08
1980	-0.31	-0.31	**2004**	0.27	0.27
1981	0.14	0.14	**2005**	-0.09	-0.09
1982	-0.17	-0.17	**2006**	0.00	0.00
1983	-0.16	-0.16	**2007**	-0.61	-0.61
1984	0.27	0.27	**2008**	0.27	0.27
1985	-0.10	-0.10	**2009**	0.19	0.19
1986	-0.22	-0.22	**2010**	-0.09	-0.09
1987	-0.02	-0.02	**2011**	0.09	0.09
1988	0.11	0.11	**2012**	0.03	0.03
1989	-0.11	-0.11	**2013**	0.07	0.07
1990	0.05	0.05	**2014**	0.00	0.00
1991	-0.06	-0.06	**2015**	0.03	0.03
1992	-0.03	-0.03	**2016**	0.34	0.34
1993	-0.09	-0.09			
			Total	-1.43	-1.43

Table 1

Max-Mean Approach

Max-Mean

For each stock, instead of selecting the median of the daily returns across the trading days in a year, we try to use mean as the representation of the annual performance. Unfortunately, this strategy does not have obvious better performance than the traditional Max-Median method. Its maximum portfolio gain is almost the same as Max-Median's maximum portfolio gain, around 30%. Its maximum portfolio loss is almost the same as Max-Median's maximum portfolio loss, around 60%. The result has shown that the Max-Mean approach is not necessarily better than the Max-Median approach.

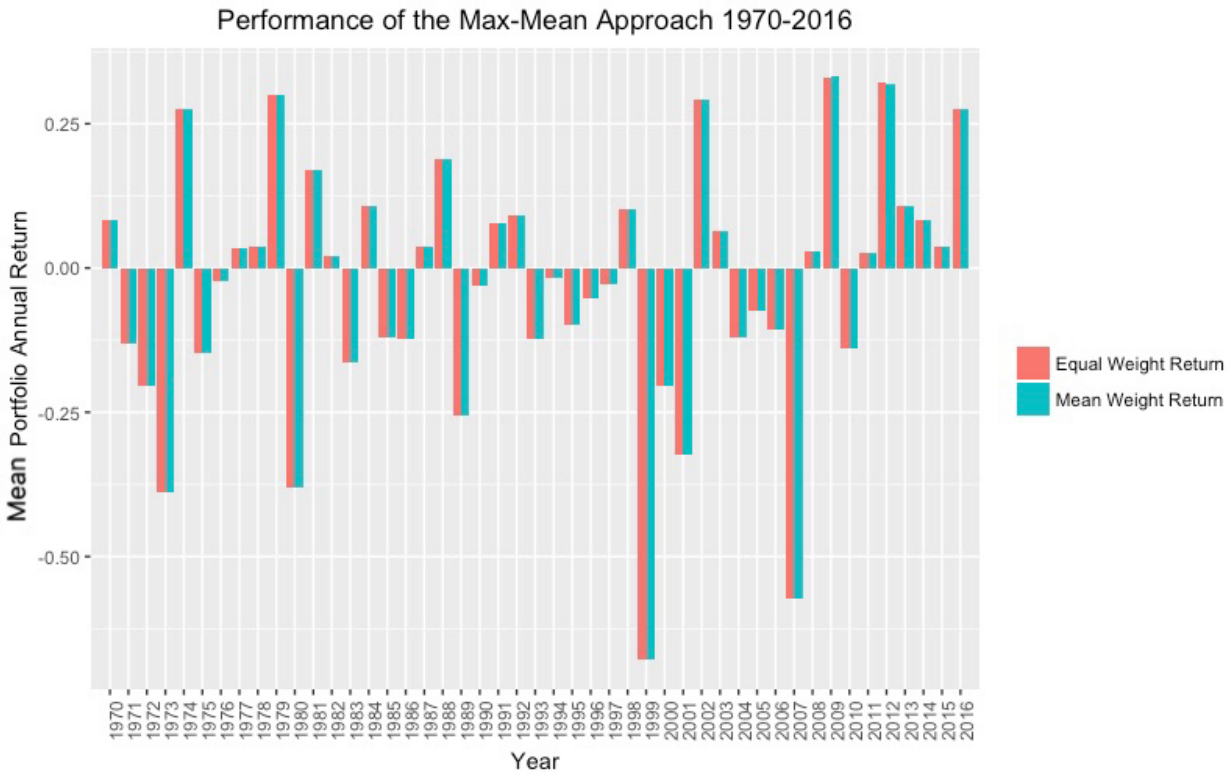

Figure 2

Table 2 shows the returns of Max-Mean portfolio from 1970 – 2016.

Year	Equal-weighted	Mean-weighted	Year	Equal-weighted	Mean-weighted
1970	0.08	0.08	**1994**	-0.02	-0.02
1971	-0.13	-0.13	**1995**	-0.10	-0.10
1972	-0.20	-0.20	**1996**	-0.05	-0.05
1973	-0.39	-0.39	**1997**	-0.03	-0.03
1974	0.28	0.28	**1998**	0.10	0.10
1975	-0.15	-0.15	**1999**	-0.68	-0.68
1976	-0.02	-0.02	**2000**	-0.21	-0.21
1977	0.03	0.03	**2001**	-0.32	-0.32
1978	0.04	0.04	**2002**	0.29	0.29
1979	0.30	0.30	**2003**	0.06	0.06
1980	-0.38	-0.38	**2004**	-0.12	-0.12
1981	0.17	0.17	**2005**	-0.08	-0.08
1982	0.02	0.02	**2006**	-0.11	-0.11
1983	-0.16	-0.16	**2007**	-0.57	-0.57
1984	0.11	0.11	**2008**	0.03	0.03
1985	-0.12	-0.12	**2009**	0.33	0.33
1986	-0.12	-0.12	**2010**	-0.14	-0.14
1987	0.04	0.04	**2011**	0.03	0.03
1988	0.19	0.19	**2012**	0.32	0.32
1989	-0.26	-0.26	**2013**	0.11	0.11
1990	-0.03	-0.03	**2014**	0.08	0.08
1991	0.08	0.08	**2015**	0.04	0.04
1992	0.09	0.09	**2016**	0.28	0.28
1993	-0.12	-0.12			
			Total	-1.42	-1.42

Table 2

Max-Mode Approach

Max-Mode

For each stock, instead of selecting the median of the daily returns across the trading days in a year, we try to use mode as the representation of the annual performance. Fortunately, this strategy performs much better than the traditional Max-Median method, increasing maximum portfolio gain from 30% to almost 60% and decreasing maximum portfolio loss from 60% to 40%. The result has shown that the mode may be more informative of an individual security's annual performance than the median.

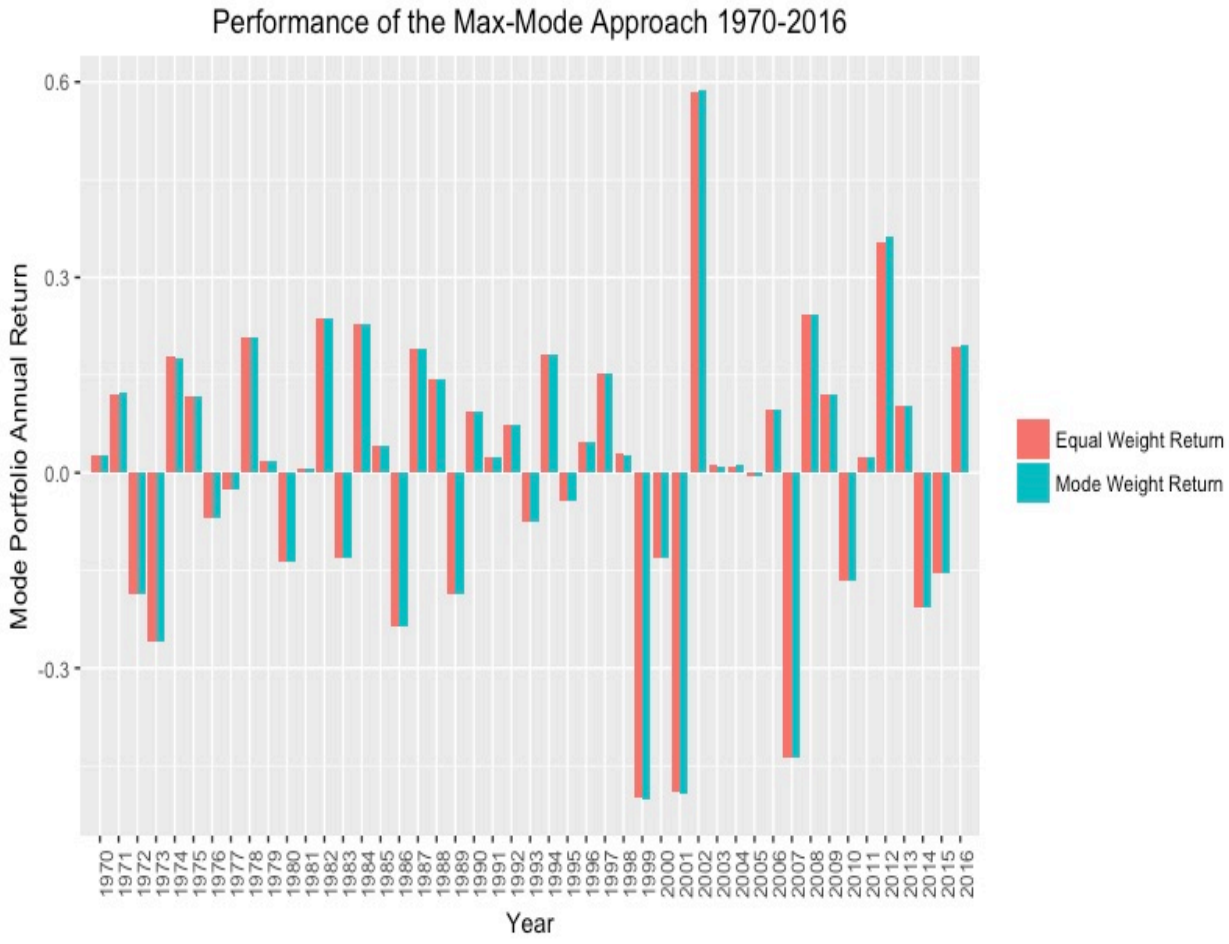

Figure 3

Table 3 shows the returns of the Max-Mode portfolio from 1970-2016.

Year	Equal-weighted	Mode-weighted	Year	Equal-weighted	Mode-weighted
1970	0.03	0.03	**1994**	0.18	0.18
1971	0.12	0.12	**1995**	-0.04	-0.04
1972	-0.19	-0.19	**1996**	0.05	0.05
1973	-0.26	-0.26	**1997**	0.15	0.15
1974	0.18	0.18	**1998**	0.03	0.03
1975	0.12	0.12	**1999**	-0.50	-0.50
1976	-0.07	-0.07	**2000**	-0.13	-0.13
1977	-0.03	-0.03	**2001**	-0.49	-0.49
1978	0.21	0.21	**2002**	0.58	0.59
1979	0.02	0.02	**2003**	0.01	0.01
1980	-0.14	-0.14	**2004**	0.01	0.01
1981	0.01	0.01	**2005**	-0.01	-0.01
1982	0.24	0.24	**2006**	0.10	0.10
1983	-0.13	-0.13	**2007**	-0.44	-0.44
1984	0.23	0.23	**2008**	0.24	0.24
1985	0.04	0.04	**2009**	0.12	0.12
1986	-0.24	-0.24	**2010**	-0.17	-0.17
1987	0.19	0.19	**2011**	0.03	0.02
1988	0.14	0.14	**2012**	0.35	0.36
1989	-0.19	-0.19	**2013**	0.10	0.10
1990	0.09	0.09	**2014**	-0.21	-0.21
1991	0.02	0.02	**2015**	-0.15	-0.15
1992	0.07	0.07	**2016**	0.19	0.19
1993	-0.08	-0.08			
			Total	0.42	0.42

Table 3

Max-Kernel Approach

Kernel

Inspired by the kernel smoothing method in time series forecasting, we repeatedly use modified Daniel kernel, a weight generating function with the feature of putting half weights at the endpoints, to the daily return sequence of each individual stock. It not only includes more information than just using median of the sequence but also assigns heavier weights to the important numbers (daily returns around the median) and fewer weights to the extreme values. The nature of Daniel kernel determines that it can only be applied to the sequence with odd length, so we truncate the original data by deleting the smallest return for even-length sequence and the stock traded less than 4 days since the information is not deterministic. Unfortunately, the resulting portfolio does not outperform the Max-Median based one, but the insight is that even it fails on Daniel kernel, the others might work and it is worth the investigation.

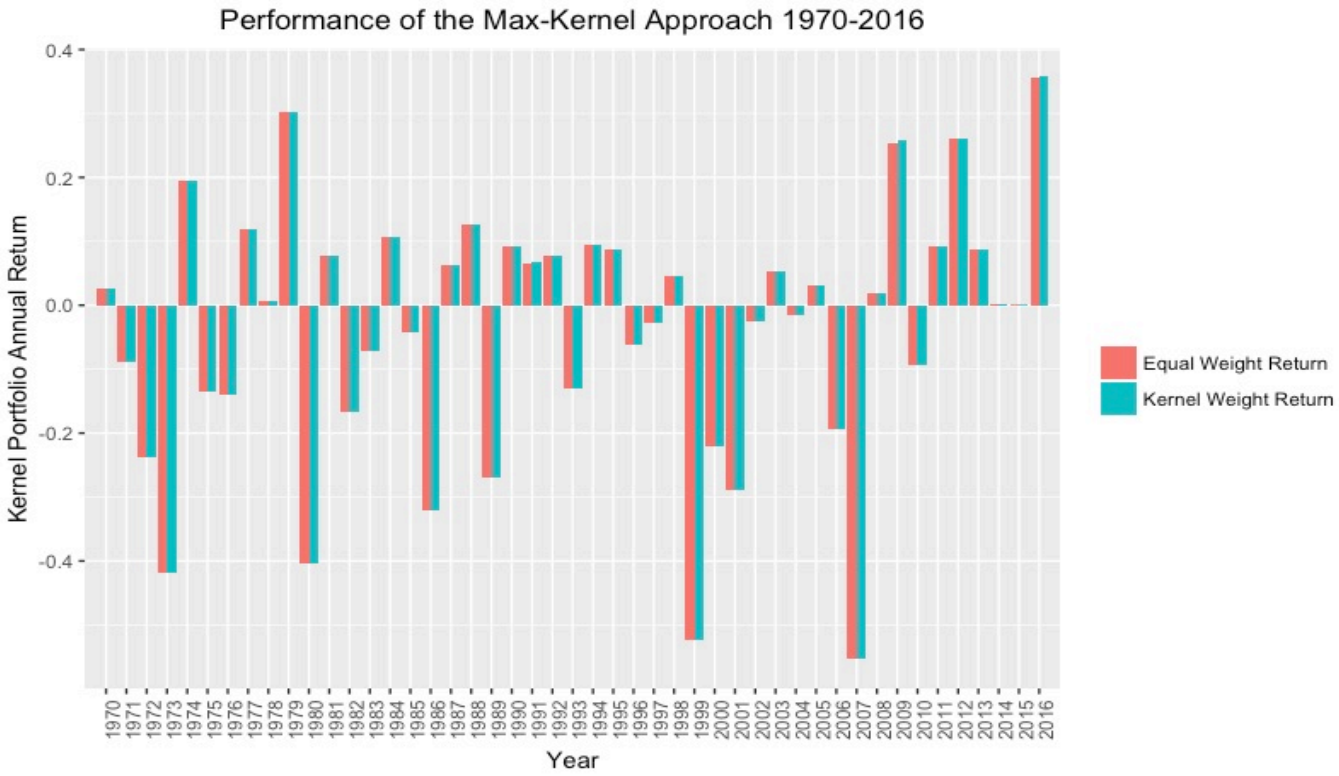

Figure 4

Table 4 shows the returns of Max-Kernel Approach from 1970-2016.

Year	Equal-weighted	Kernel-weighted	Year	Equal-weighted	Kernel-weighted
1970	0.03	0.03	**1994**	0.09	0.09
1971	-0.09	-0.09	**1995**	0.09	0.09
1972	-0.24	-0.24	**1996**	-0.06	-0.06
1973	-0.42	-0.42	**1997**	-0.03	-0.03
1974	0.20	0.20	**1998**	0.05	0.05
1975	-0.14	-0.14	**1999**	-0.52	-0.52
1976	-0.14	-0.14	**2000**	-0.22	-0.22
1977	0.12	0.12	**2001**	-0.29	-0.29
1978	0.01	0.01	**2002**	-0.02	-0.02
1979	0.30	0.30	**2003**	0.05	0.05
1980	-0.40	-0.40	**2004**	-0.01	-0.01
1981	0.08	0.08	**2005**	0.03	0.03
1982	-0.17	-0.17	**2006**	-0.19	-0.19
1983	-0.07	-0.07	**2007**	-0.55	-0.55
1984	0.11	0.11	**2008**	0.02	0.02
1985	-0.04	-0.04	**2009**	0.25	0.26
1986	-0.32	-0.32	**2010**	-0.09	-0.09
1987	0.06	0.06	**2011**	0.09	0.09
1988	0.13	0.13	**2012**	0.26	0.26
1989	-0.27	-0.27	**2013**	0.09	0.09
1990	0.09	0.09	**2014**	0.00	0.00
1991	0.07	0.07	**2015**	0.00	0.00
1992	0.08	0.08	**2016**	0.36	0.36
1993	-0.13	-0.13			
			Total	-1.79	-1.78

Table 4

Conclusion

In this report, we query, collect, and filter data from CRSP-Compustat database, and reproduce the result of portfolio management by using the Max-Median Rule. We also manage to improve the results by modifying the algorithm to select the 20 stocks with max daily ratio mean, mode, and kernel, instead of max daily ratio median.

The Max-Kernel approach combines the ideas from the Max-Median Rule and the Max-Mean approach. Since the two original approaches do not perform well, the Max-Kernel approach also underperforms the S&P500 Index. However, the result we got from the Max-Mode approach outperforms that of the Max-Median rule, enlightening us to improve the Max-Median Rule.

REFERENCE

Services, W. (2018). *CRSP Daily Stock*. [online] Wrds-web.wharton.upenn.edu. Retrieved from: https://wrdsweb.wharton.upenn.edu/wrds//ds/crsp/stock_a/dsf.cfm?queryID=855937. Accessed Apr 06, 2018.

Thompson, James R., Baggett, L. Scott. “Everyman’s MaxMedian Rule for Portfolio Management.” Rice University, 2009.

Code Appendix

```
### STAT486 Mini Project 4: Max-Median Rule
library(ggplot2)
path.data = "/Users/gyuanrui/Desktop/STAT/486/project_4/data/"
path.file = "/Users/gyuanrui/Desktop/STAT/486/project_4/"

mode <- function(x) {
  ux <- unique(x)
  ux[which.max(tabulate(match(x, ux)))]
}

yr_mmreturn <- data.frame(matrix(ncol = 3, nrow = 0))
colnames(yr_mmreturn) <- c("year", "ew_r", "mw_r")
yr_return <- data.frame(matrix(ncol = 3, nrow = 0))
colnames(yr_return) <- c("year", "return", "type")
yrcount <- 1

for (year in c(1970:2016)) {
  yrdata <- read.table(paste(path.data, year, ".csv", sep = ""),
header = TRUE, sep = ",")
  yrdata <- yrdata[c("PERMNO", "date", "PRC")]
  yrdata <- na.omit(yrdata)
  yrdata$year <- format(as.Date(as.character(yrdata$date),
format="%Y%m%d"),"%Y")
  yrlist <- data.frame(matrix(ncol = 4, nrow = 0))
  colnames(yrlist) <- c("PERMNO", "median", "initial", "ending")
  cocount <- 1
  for (company in unique(yrdata$PERMNO)) {
    cur_yr <- yrdata[yrdata$PERMNO == company, ]
    cur_yr <- yrdata[yrdata$PERMNO == company & yrdata$year ==
year, ]
    cur_yr <- cur_yr[order(as.Date(as.character(cur_yr$date),
format = "%Y%m%d")), ]
    next_yr <- yrdata[yrdata$PERMNO == company & yrdata$year ==
year + 1, ]
    next_yr <- next_yr[order(as.Date(as.character(next_yr$date),
format = "%Y%m%d")), ]
    if (nrow(cur_yr) != 0 && nrow(next_yr) != 0) {
      ratio <- c()
      for (i in seq(2, nrow(cur_yr))) {
        ratio <- c(ratio, as.numeric(cur_yr$PRC[i]) /
as.numeric(cur_yr$PRC[i-1]))
      }
      m <- median(ratio)
      initial <- next_yr$PRC[1]
```

```
      ending <- next_yr$PRC[nrow(next_yr)]
      yrlist[cocount, ] <- c(company, m, initial, ending)
      cocount <- cocount + 1
    }
  }
  top20 <- yrlist[order(- yrlist$median), ][1:20, ]
  ew_r <- (sum(top20$ending) - sum(top20$initial)) /
sum(top20$initial)
  mw <- top20$median / sum(top20$median)
  mw_r <- (sum(top20$ending * mw) - sum(top20$initial * mw)) /
sum(top20$initial * mw)
  yr_mmreturn[yrcount, ] <- c(year, ew_r, mw_r)
  yr_return[yrcount * 2 - 1, ] <- c(year, ew_r, "equal_weighted")
  yr_return[yrcount * 2, ] <- c(year, mw_r, "median_weighted")
  yrcount <- yrcount + 1
}

View(yr_mmreturn)

write.csv(yr_mmreturn, file = paste(path.file, "kernel.csv", sep
= ""), row.names = FALSE, na="")
write.csv(yr_return, file = paste(path.file, "returns8.csv", sep
= ""), row.names = FALSE, na="")


yr_return <- read.table(paste(path.file, "median2.csv", sep =
""), header = TRUE, sep = ",")
type <- yr_return$type
yr_return <- data.frame(sapply(yr_return, function(x)
as.numeric(as.character(x))))
yr_return$type <- type
sp500 <- read.table(paste(path.data, "s&p500.csv", sep = ""),
header = TRUE, sep = ",")
i <- 1
for (year in sp500$X) {
  if (year >= 1970 && year <= 2016) {
    yr_return[nrow(yr_return) + 1, ] = c(year, sp500$ewx[i],
"sp500")
    print(c(year, sp500$ewd[i], "sp500"))
  }
  i <- i + 1
}
yr_return
type <- yr_return$type
yr_return <- data.frame(sapply(yr_return, function(x)
as.numeric(as.character(x))))
yr_return$type <- type
yr_return
```

```
ggplot(yr_return, aes(x = year, y = return, fill = type)) +
  geom_bar(stat = "identity", position = "dodge") +
  ggtitle("Performance of the Max-Median Rule 1970-2016") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1),
        plot.title = element_text(hjust = 0.5)) +
  labs(x = "Year", y = "Median Portfolio Annual Return") +
  scale_x_discrete(limits = seq(1970, 2016, 1)) +
  scale_fill_discrete(name = "", labels = c("Equal Weight
Return", "Median Weight Return", "S&P500 Return"))
```