Portfolio performance

Jesse Keränen

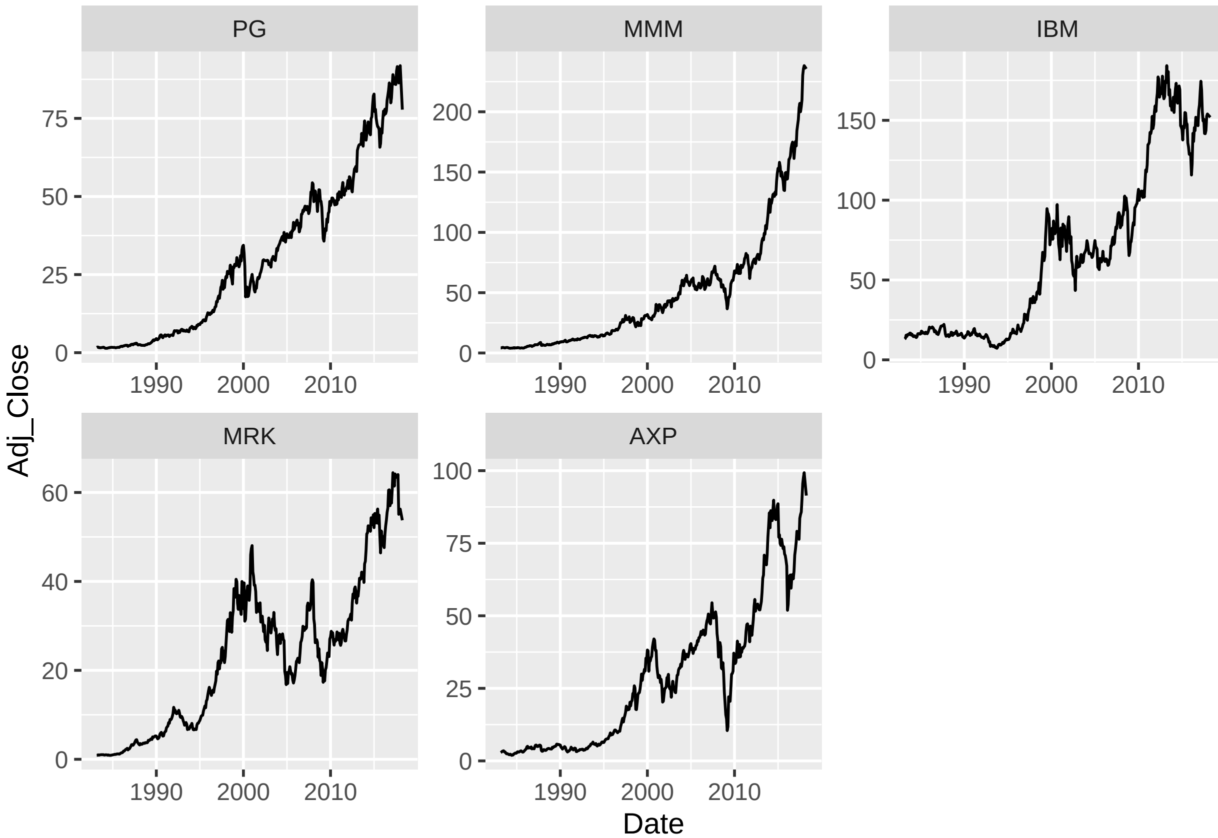
2/4/2022

Prologue

Quite as important as to be able to construct a efficient portfolio is the capability to analyze historical performance of portfolio. Many times people just look at the return of an asset, but that's just the other side of the coin. It is merely as important to look how much risk was taken to achieve the returns. This time I change stocks that I deal with. I have chosen five stocks that have remained in Dow Jones Industrial Average-index more than 35 years. These stocks are The Protector & Gamble Company (PG), 3M Company (MMM), International Business Machines Company(IBM), Merck & Co., Inc (MRK) and American Express Company (AXP). I have data from the year 1983 to 2018. What I am going to do is to every year calculate minimum variance portfolio from these stock and compare performance of my portfolio to Dow Jones Industrial-index. I kind of simulate actions of hypothetical investor. One could argue that my experiment suffers from survivor bias and that is true. I Have chosen companies that have remained among the biggest companies in United States for decades and that pretty much indicates they have been really successful. On the other hand turnover within Dow Jones Industrial-index isn't too high. From companies that are included in index today approximately half have been there more than 20 years. So i guess that, at least in history, investor haven't had to bee too lucky to be able to pick stock that remain moderately long period of time in the index.

```
library(ggplot2)
library(Quandl)
library(tidyverse)
library(data.table)
library(zoo)
library("xlsx")

Quandl.api_key("bx1qdehfWXg6SNKnicQC")

names <- c("PG", "MMM", "IBM", "MRK", "AXP")
weights <-c(0.3, 0.1, 0.2, 0.2, 0.2)

# For monthly data use collapse = "monthly"
prices <- as.data.table(Quandl(c("WIKI/PG.11","WIKI/MMM.11",
        "WIKI/IBM.11", "WIKI/MRK.11", "WIKI/AXP.11"), start_date
        = "1983-01-01", end_date = "2021-12-31", collapse = "monthly"))

prices <- na.omit(prices)

colnames(prices) <- c("Date", names)

# Tidying data table
prices <- melt(prices, id.vars = "Date", measure.vars = names, variable.name
               = "Company", value.name = "Adj_Close")
```

```r
prices[, "Returns" := Adj_Close/shift(Adj_Close) -1, by = Company]
prices <- drop_na(prices)

# Development of the stock prices
ggplot(prices, aes(Date, Adj_Close)) + geom_line() + facet_wrap(~Company, scales = "free")
```

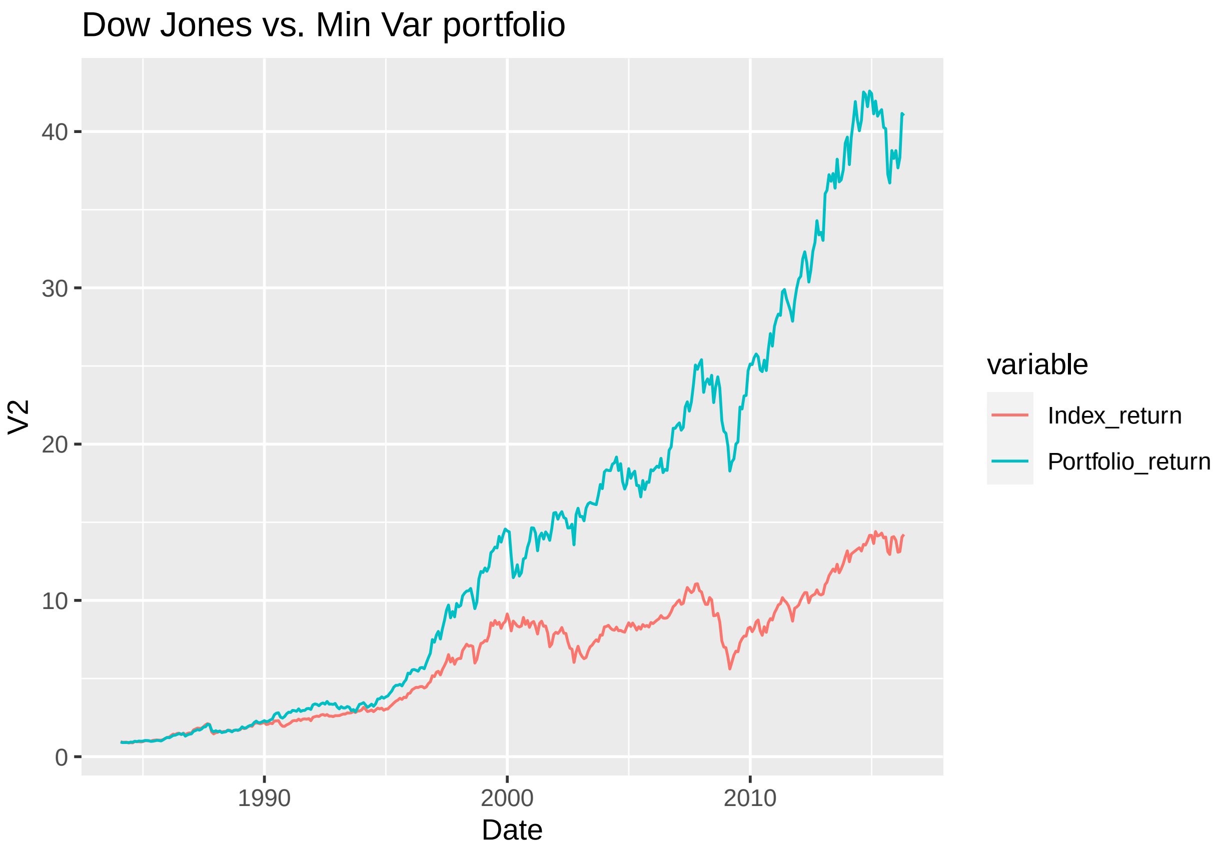


Portfolio construction

Idea is to adjust weights within the portfolio January of each year to construct minimum variance portfolio. Minimum variance portfolio comes handy since we don't have to express opinion on expected returns of underlying stocks. Variances and etc. are calculated using historical data from last five years. It would make experiment more realistic to add a cost for rebalancing the portfolio, but let's leave that out of this experiment at least for the first time. It probably will be easiest to write a function that calculates optimal weights from data from last five years. Then we can call this function every year. Since we want to form minimum variance portfolio our optimization problem comes pretty easy. For more precise details see Markowitz-file. We don't restrict short selling.

```r
optimize <- function(dt, date){
  data <- dt[year(Date) < date & year(Date) > year(date) - 6]
  cov_matrix <- cov(as.data.table(split(data[,4] ,data$Company)))
  inverse <- solve(cov_matrix)
  ones <- matrix(rep(1, length(names)))
  c <- t(ones) %*% inverse %*% ones
  weights <- as.numeric(1/c) * (inverse %*% ones)
```

```
  return(data.table(Date = date, Weights = weights, Company =
                    row.names(weights)))
}

januarys <- prices[month(Date) == 1, Date]
januarys <- januarys[5:length(januarys)]

opt_weights <- rbindlist(lapply(januarys, function(t){optimize(prices , t)}))
opt_weights <- unique(opt_weights)


prices <- merge(prices, opt_weights, by = c("Company", "Date"), all.x = T)

prices <- na.locf(prices, na.rm = F)

portfolio_returns <- prices[, sum(Returns*Weights.V1), by = Date]
portfolio_returns <- drop_na(portfolio_returns)
# TODO remove data from 2018

colnames(portfolio_returns)[2] <- "Portfolio_return"

dow_jones <- as.data.table(Quandl(c("BCB/UDJIAD1"), start_date
        = "1983-01-01", end_date = "2021-12-31", collapse = "monthly"))
dow_jones <- dow_jones[order(Date)]
dow_jones[, Change := Value/shift(Value)-1]


portfolio_returns_melt <- merge(dow_jones, portfolio_returns, by = "Date")
colnames(portfolio_returns_melt) <- c("Date", "Index_value", "Index_return", "Portfolio_return")

portfolio_returns_melt <- melt(portfolio_returns_melt, id.vars = "Date", measure.vars = c("Index_return

# Seems like our portfolio has remarkably outperformed index
ggplot(portfolio_returns_melt[, .(Date, cumprod(1+value)), by = variable], aes(Date, V2, color = variabl
```




```r
portfolio_returns_melt[, .(sd(value), mean(value)), by = variable]
```

```
##            variable         V1          V2
## 1:     Index_return 0.04329714 0.007818154
## 2: Portfolio_return 0.04231149 0.010510105
```

Portfolio performance

So it seems that optimization has gained us profit. Undeniably return of our portfolio has been remarkably higher than index on average. Let's see if we have been able to generate profit without being exposed to higher amount of risk. We have to remember that our objective was to minimize variance not as high return as possible. Next look what some of the well know performance measurements tell us about performance of our portfolio. I will calculate 4 different measurements. Sharpe ratio tells us ratio between standard deviation of our portfolio and excess returns. Treynor's ratio is similar but this time excess returns are compared only to systematic risk (beta) of the investment. Jensen's alpha measures if there is constant excess returns compared to estimates of Cap-model. Also Omega ratio will be discussed. It tells us for given threshold value weighted average of loss profit ratio.

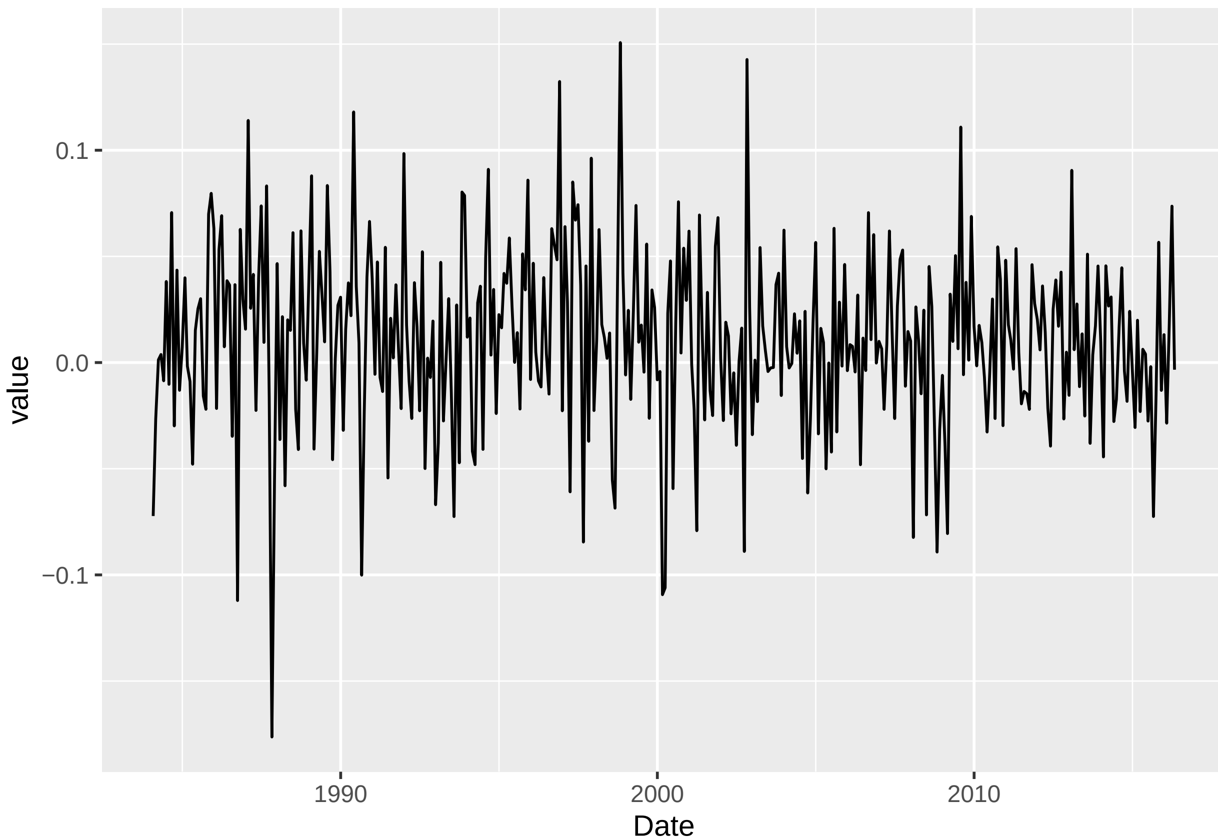
Sharpe ratio

Sharpe ratio is very much used and easily understandable measurement. It measures ratio between investments excess returns and it's standard deviation. I use monthly data, but annualize estimates before calculating the ratio.

$$Sharpe = \frac{\overline{r}_x - r_f}{SD(r_x)}$$

Where $\overline{r}_x$ is expected return of the investment, r_f is risk free rate and $SD(r_x)$ is the standard deviation of the returns of the investment.

```
ggplot(portfolio_returns_melt[variable == "Portfolio_return"], aes(Date, value)) + geom_line()
```




```
factors <- as.data.table(read.table("extdata/FF5_factors.txt",header=T))
factors <- factors/100
end_dates <- seq(as.Date("1963-08-01"),length=702,by="months")-1
factors[, Date := as.Date(end_dates)]

first_date <- min(portfolio_returns$Date)
last_date <- max(portfolio_returns$Date)
factors <- factors[Date >= first_date & Date <= last_date]

portfolio_returns <- merge(portfolio_returns, factors, by = "Date")
portfolio_returns[, Portfolio_return := Portfolio_return - RF]

rf <- tail(portfolio_returns$RF, n=1)

sharpe <- ((mean(portfolio_returns$Portfolio_return)-rf)*12)/(sd(portfolio_returns$Portfolio_return)*sq
sharpe
```

```
## [1] 0.5262496
```

Sharpe ratio of our investment haven't been too good. Of course one can not say too much about just one number. It would make more sense to calculate Sharpe ratios for other investments as well and compare our Sharpe to those. Nevertheless I have read that as a rule of thumb investor can consider Sharpe ratio to be good when it is bigger than one.

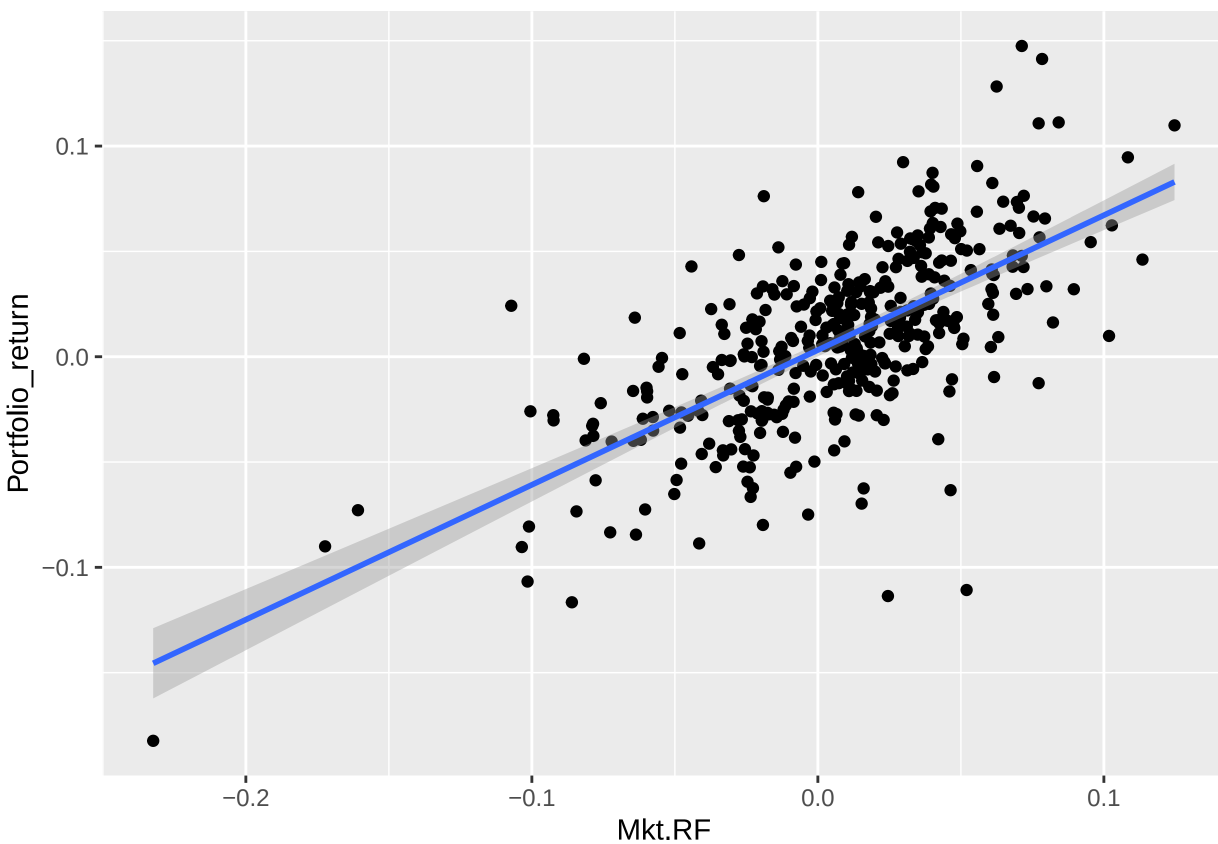
Treynor ratio

I will use data from Kenneth French's data base for market factors. In recent years risk free rates have been close to zero, but this might not have been the case in the 80's. The dates are formatted differently in French's data sets. So I will create new array of date similar to our previous data tables and set this vector to dates in the factor data table. This will make our work easier in the future.

$$Treynor = \frac{\overline{r}_x - r_f}{\beta_x}$$

```
ggplot(portfolio_returns, aes(Mkt.RF, Portfolio_return)) + geom_point() + geom_smooth(method = "lm")
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
capm <- lm(data = portfolio_returns, Portfolio_return ~ Mkt.RF)
beta <- unname(capm$coefficients[2])
beta
```

```
## [1] 0.6399599
```

```r
treynor <- ((mean(portfolio_returns$Portfolio_return)-rf)*12)/beta
treynor
```

```
## [1] 0.1180866
```

First of all beta of our portfolio has been about 0.64. That means that our investment has been less volatile than markets on average. This is of course pretty obvious since we minimized variance of our portfolio. Treynor ratio is approximately 0.12. Again similar as with the Sharpe ratio we can't draw meaningful conclusions from just one Treynor ratio. It is best used in ranking portfolios. Investment with higher Treynor ratio is better than one with lower, but it is hard to say how much better.

Jensen's alpha

Jensen's alpha solves one problem we had with our previous measurements. It can tell us also how much better one investment performed than another. With R calculating alpha is pretty easy using linear regression and excess returns. Following equation gives us the alpha:

$$\alpha = (r_i - r_f) - \beta(r_m - r_f)$$

Where r_m is the market return.

```r
alpha <- capm$coefficients[1]
alpha
```

```
## (Intercept)
##  0.00317165
```

```r
# Let's takes a closer look on statistical significance
summary(capm)
```

```
##
## Call:
## lm(formula = Portfolio_return ~ Mkt.RF, data = portfolio_returns)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.147207 -0.019090 -0.000647  0.019330  0.098716
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.003172   0.001537   2.063   0.0397 *
## Mkt.RF      0.639960   0.034923  18.325   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03072 on 407 degrees of freedom
## Multiple R-squared:  0.4521, Adjusted R-squared:  0.4507
## F-statistic: 335.8 on 1 and 407 DF,  p-value: < 2.2e-16
```

Alpha of our portfolio is positive. This is a good thing. We have been able to generate profits that cannot be explained by our exposure to market risk. It is even statistically significant. Maybe not as strongly as we would have hoped, but at least in the five percent confidence level. As portfolio managers we probably would be quite happy to these results. Nevertheless, as customers of the portfolio manager we again probably would want to take another look. Some studies, most notably Fama and French (1993), have show that there are some persistent risk factors in the market other than market risk. For example firms with small market capitalization tend to have higher returns than firms with high market capitalization. Other example is that so called value firms, firms with high book-to-market ratio, tend to have higher returns than firms with low book-to-market ratio. There isn't too good explanation yet why firms with this kind of characteristics are considered riskier by the markets. As a customer we might want to look if the excess returns were in reality compensation for these risks. Fama French model is discussed in more detail in FamaFrench-file.

```
FF3 <- lm(data=portfolio_returns, Portfolio_return ~ Mkt.RF + SMB + HML)
FF5 <- lm(data=portfolio_returns, Portfolio_return ~ Mkt.RF + SMB + HML + RMW + CMA)

summary(FF3)
```

```
##
## Call:
## lm(formula = Portfolio_return ~ Mkt.RF + SMB + HML, data = portfolio_returns)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.196616 -0.016922 -0.001579  0.017551  0.096580
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.003258   0.001478   2.205    0.028 *
## Mkt.RF       0.676852   0.034739  19.484  < 2e-16 ***
## SMB         -0.329735   0.050758  -6.496 2.42e-10 ***
## HML         -0.040621   0.051632  -0.787    0.432
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02931 on 405 degrees of freedom
## Multiple R-squared:  0.5038, Adjusted R-squared:  0.5002
## F-statistic: 137.1 on 3 and 405 DF,  p-value: < 2.2e-16
```

```
summary(FF5)
```

```
##
## Call:
## lm(formula = Portfolio_return ~ Mkt.RF + SMB + HML + RMW + CMA,
##     data = portfolio_returns)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.196258 -0.016607 -0.000393  0.016594  0.082589
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.0003588  0.0014490   0.248    0.805
## Mkt.RF       0.7880579  0.0361621  21.792  < 2e-16 ***
```

```
## SMB         -0.2368929  0.0524357  -4.518 8.22e-06 ***
## HML         -0.3354513  0.0660782  -5.077 5.88e-07 ***
## RMW          0.3709849  0.0672530   5.516 6.20e-08 ***
## CMA          0.5303373  0.0985041   5.384 1.24e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02762 on 403 degrees of freedom
## Multiple R-squared:  0.5617, Adjusted R-squared:  0.5562
## F-statistic: 103.3 on 5 and 403 DF,  p-value: < 2.2e-16
```

Alpha remains more or less unchanged and as significant with the three factor model as with the Cap model. However, when we introduce two additional variables, namely profitability and investment, alpha isn't significant anymore. Interestingly all of our factors become highly significant.