Portfolio Management on NVDA, CMCSA and EA stocks

Paul Muriithi

```r
# loading required libraries
library(tidyquant)
library(tidyverse)
library(broom)
library(knitr)
library(kableExtra)
library(ggplot2)
library(plotly)
library(gridExtra)
```

1. Importing stocks data

```r
# Fetch selected stocks using Tiingo
stocks <- tq_get(c("NVDA", "CMCSA", "EA"), get = "stock.prices", from = "2000-01-01", to = "2022-06-18"
    select(symbol, date, adjusted)


# Display first 6 rows in a table
head(stocks, n = 6) %>%
    kable(caption = "Selected stocks data")
```

Table 1: Selected stocks data

symbol	date	adjusted
NVDA	2000-01-03	0.0894252
NVDA	2000-01-04	0.0870375
NVDA	2000-01-05	0.0841721
NVDA	2000-01-06	0.0786796
NVDA	2000-01-07	0.0799936
NVDA	2000-01-10	0.0826197

2. The Analysis

2.1. Plot prices over time.

```r
# time series plot for the three stocks
NVDA = stocks[stocks$symbol == "NVDA", ]
CMCSA = stocks[stocks$symbol == "CMCSA", ]
EA = stocks[stocks$symbol == "EA", ]
plot1 <- ggplot(NVDA, aes(x = date, y = adjusted)) + geom_line(color = "darkblue") + labs(title = "NVDA
plot2 <- ggplot(CMCSA, aes(x = date, y = adjusted)) + geom_line(color = "orange") + labs(title = "CMCSA
```

```r
plot3 <- ggplot(EA, aes(x = date, y = adjusted)) + geom_line(color = "tomato") + labs(title = "EA prices
grid.arrange(plot1, plot2, plot3, ncol = 3)
```

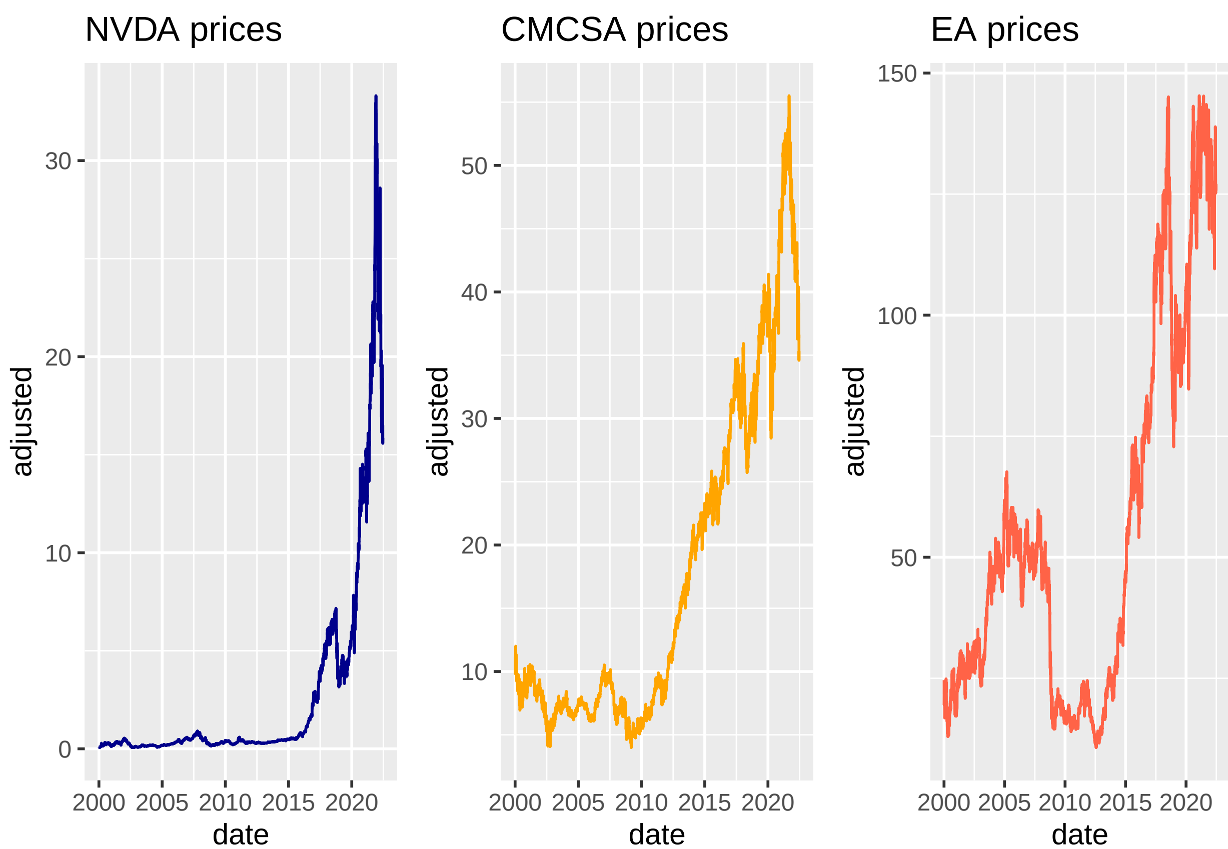

Figure 1: Time series of stocks prices

NVDA: The stock prices were very low and stationary from 2000 to around 2016 after when they increased rapidly with cyclic patterns. The prices were at their maximum in 2021, but could not maintain that price where its followed by a decreasing trend.

CMCSA: The stock prices were low from 2000 to 2010. Since 2010, the prices have been on sharp increasing trend with seasonal properties. This stock reaches its peak in 2021, falling immediately thereafter.

EA: EA stock prices started in 2000 with an upward trend where they rose for 5 yrs to 2005, then started to depreciate. At around 2008, EA stock prices fell sharply to where they were in 2000 and remained so with some seasonal rise and falls up to mid-2012. The prices then skyrocketed, reaching their peak in 2018 where they immediately falls again in 2019 and 2020, but rises in 2021.

2.2.

Given the formula:

$$r_t = 100 * \ln\left(\frac{P_t}{P_{t-1}}\right)$$

```r
# percentage returns for each stock
stocks <- stocks %>%
    group_by(symbol) %>%
```

```r
    mutate(per_returns = 100 * (log(adjusted) - log(lag(adjusted)))) %>%
    ungroup() %>%
    drop_na()

# subset percentage returns for each stock
NVDA_RT = stocks[stocks$symbol == "NVDA", ]
CMCSA_RT = stocks[stocks$symbol == "CMCSA", ]
EA_RT = stocks[stocks$symbol == "EA", ]
# rendering plots fro the percentage returns
plt1 <- ggplot(CMCSA_RT, aes(x = date, y = per_returns)) + geom_line(color = "darkblue") +
    labs(title = "NVDA percentage returns")
plt2 <- ggplot(CMCSA_RT, aes(x = date, y = per_returns)) + geom_line(color = "orange") + labs(title = "
plt3 <- ggplot(EA_RT, aes(x = date, y = per_returns)) + geom_line(color = "tomato") + labs(title = "EA
plts <- grid.arrange(plt1, plt2, plt3, ncol = 2)
```

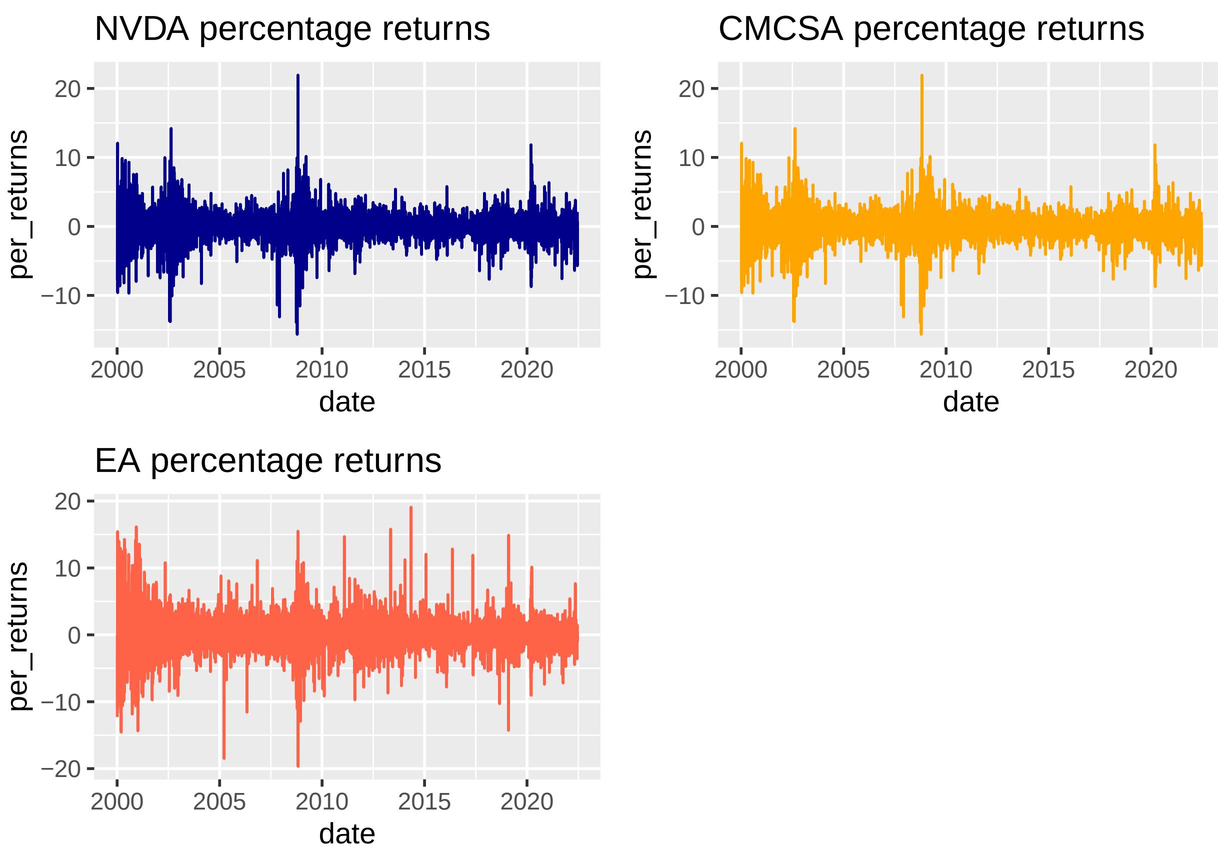

Figure 2: Time series of percentage returns

```r
plts
```

```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z     cells    name           grob
## 1 1 (1-1,1-1) arrange gtable[layout]
## 2 2 (1-1,2-2) arrange gtable[layout]
## 3 3 (2-2,1-1) arrange gtable[layout]
```

2.3

To find the number of bins, consider the following formula:

$$bins = (Max - Min)/h$$

where

$$h = 2 \times IQR \times n^{-1/3}$$

```r
# histogram for each stock return series.
hist1 <- ggplot(NVDA_RT, aes(x = per_returns)) + geom_histogram(bins = 151, fill = "darkblue") +
    labs(title = "NVDA returns Histogram")
hist2 <- ggplot(CMCSA_RT, aes(x = per_returns)) + geom_histogram(bins = 96, fill = "orange") +
    labs(title = "CMCSA returns Histogram")
hist3 <- ggplot(EA_RT, aes(x = per_returns)) + geom_histogram(bins = 91, fill = "tomato") +
    labs(title = "EA returns Histogram")
hists <- grid.arrange(hist1, hist2, hist3, ncol = 2)
```

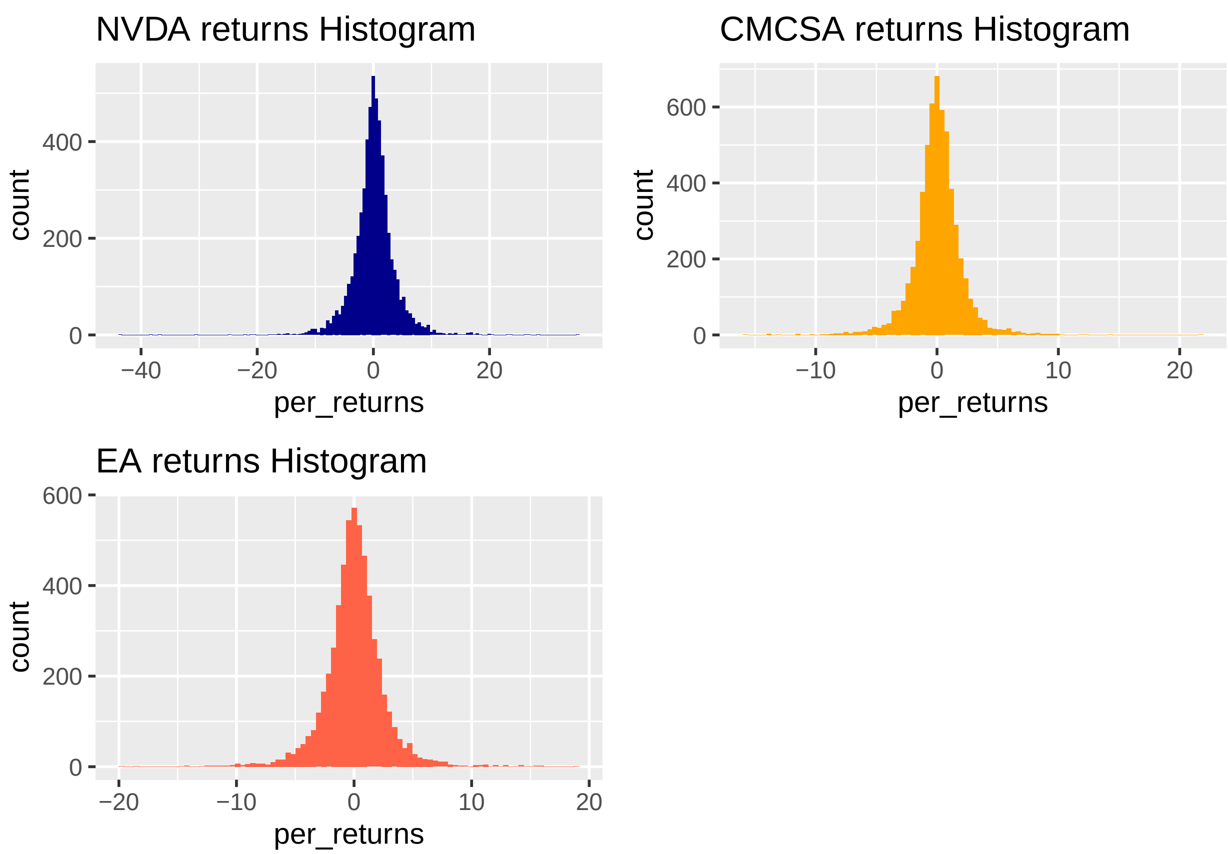

Figure 3: Histograms of percentage returns

```r
hists
```

```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z     cells    name           grob
## 1 1 (1-1,1-1) arrange gtable[layout]
## 2 2 (1-1,2-2) arrange gtable[layout]
```

```
## 3 3 (2-2,1-1) arrange gtable[layout]
```

The three histograms indicates that the returns of the three stocks(NDVA, CMCSA, and EA) are normally distributed with sharp apex indicating highly positive kurtosis, hence leptokurtic. In all three stocks, most of the stocks returns cluster around zero, since more returns are above the normal distribution and less values distributed in the tail.

2.4

```r
# Summary for adjusted prices
tab1 <- stocks %>%
    group_by(symbol) %>%
    summarize(symbol = first(symbol), type = "adjusted", mean = mean(adjusted), median = median(adjuste
        variance = var(adjusted), sd = sd(adjusted), skewness = skewness(adjusted), kurtosis = kurtosis
        .groups = "drop")

tab2 <- stocks %>%
    group_by(symbol) %>%
    summarize(symbol = first(symbol), type = "returns", mean = mean(per_returns), median = median(per_r
        variance = var(per_returns), sd = sd(per_returns), skewness = skewness(per_returns),
        kurtosis = kurtosis(per_returns), .groups = "drop")

# Combine and display
summary <- bind_rows(tab1, tab2) %>%
    select(symbol, type, mean, median, variance, sd, skewness, kurtosis)

kable(summary, caption = "Table 2: Summary statistics table of returns", digits = 7)
```

Table 2: Table 2: Summary statistics table of returns

symbol	type	mean	median	variance	sd	skewness	kurtosis
CMCSA	adjusted	17.5230157	9.7002220	171.062478	13.079085	0.9754862	-0.2535478
EA	adjusted	55.2422374	46.9115143	1441.618225	37.968648	0.8662519	-0.4606879
NVDA	adjusted	2.6750335	0.3993260	29.762543	5.455506	2.9942825	9.1419180
CMCSA	returns	0.0207890	0.0341273	4.192189	2.047483	-0.0116408	8.2701270
EA	returns	0.0289470	0.0355201	6.749487	2.597977	0.1248258	6.4048950
NVDA	returns	0.0916293	0.1050301	14.276245	3.778392	-0.2535757	13.3769428

The average stock returns do not appear to be significantly different from zero(0.0204, 0.0289, and 0.0913) since more returns are above the normal distribution and lies between 2 and 4 standard deviations. NVDA stock returns the biggest return of 0.0913, with an almost 10% return on investment, but experiences the most variations.

2.5

Significant test for all stocks (t-test)

- Null Hypothesis: Average returns equals zero
- Alternative Hypothesis: Average returns not equals zero

$$H_0 : \mu = 0$$

$$H_1 : \mu \neq 0$$

$$\alpha = 0.01$$

$$t = (\frac{\hat{\mu}\sqrt{N}}{\sigma})t_{N-1}$$

Significant test for NVDA (t-test)

```r
# select returns for NVDA stock
nvda_ret <- subset(stocks, stocks$symbol == "NVDA") %>%
    select(per_returns)
# t-test for NVDA stock
nvda_test = t.test(nvda_ret$per_returns, mu = 0, conf.level = 0.99)
res <- data.frame(nvda_test$statistic, nvda_test$p.value)
names(res) <- c("t", "p-value")
res
```

```
##          t    p-value
## t 1.823014 0.06835408
```

$$t = 1.837498$$

$$p - value = 0.0661 > \ 0.05$$

Significant test for CMCSA

```r
# returns for CMCSA stock
cmcsa_ret <- subset(stocks, stocks$symbol == "CMCSA") %>%
    select(per_returns)
# t-test for CMCSA stock
cmcsa_test = t.test(cmcsa_ret$per_returns, mu = 0, conf.level = 0.99)
cmcsa_res <- data.frame(cmcsa_test$statistic, cmcsa_test$p.value)
names(cmcsa_res) <- c("t", "p-value")
cmcsa_res
```

```
##           t   p-value
## t 0.7632668 0.4453362
```

$$t = 0.7872869$$

$$p - value = 0.4311469 > \ 0.05$$

Significant Test (t-test) for EA

```r
ea_ret <- subset(stocks, stocks$symbol == "EA") %>%
    select(per_returns)  #select returns for EA stock
# t-test for EA stock
ea_test = t.test(ea_ret$per_returns, mu = 0, conf.level = 0.99)
ea_res <- data.frame(ea_test$statistic, ea_test$p.value)
names(ea_res) <- c("t", "p-value")
ea_res
```

```
##           t  p-value
## t 0.8375873 0.402298
```

$$t = 0.8405$$

$$p - value = 0.4007 > \ 0.05$$

```r
# t critical value
qt(0.01, 5648, lower.tail = F)
```

```
## [1] 2.327008
```

$$t_{critical} = 2.327$$

- summary of the three tests

```r
tests <- data.frame(Type = c("NVDA", "CMCSA", "EA"), T_stat = c(1.8375, 0.7873, 0.8405), T_crit = c(2.32
    2.327, 2.327), P_value = c(0.06612, 0.4311, 0.4007))
kable(tests, caption = "Tests of significant differences from zero")
```

Table 3: Tests of significant differences from zero

Type	T_stat	T_crit	P_value
NVDA	1.8375	2.327	0.06612
CMCSA	0.7873	2.327	0.43110
EA	0.8405	2.327	0.40070

The t-values obtained for all the three stocks are less than the critical value(2.327), therefore we fail to reject the null hypothesis and conclude that the stocks returns are not significantly different from 0 at 1% significance level. The p-values being higher than 0.01 suggest that we do not have sufficient evidence to say that the average returns are different from 0.

2.6. Significance differences of the average returns.

Testing for equality of variances

$$Let\ Variance\ NVDA\ =\ \sigma_N^2$$
$$Let\ variance\ CMCSA\ =\ \sigma_C^2$$
$$Let\ variance\ EA\ =\ \sigma_E^2$$

* **1. NVDA and CMCSA**

$$H_0 : \sigma_N^2 = \sigma_C^2$$

$$H_1 : \sigma_N^2 \neq \sigma_C^2$$

$$\alpha = 0.01 \quad \frac{\alpha}{2} = 0.005$$

The test statistic

$$F = \left(\frac{\hat{\sigma}_N^2}{\hat{\sigma}_C^2}\right)$$

$$F = \frac{14.275}{4.187} = 3.409$$

$$F_c = F_{N_G-1,N_C-1} = F_{5649-1,5649-1}$$

F critical can be obtained using r as follows:

```
qf(p = 0.005, df1 = 5649 - 1, df2 = 5649 - 1, lower.tail = FALSE)
```

```
## [1] 1.070963
```

$$F_c = 1.070963$$

$$F > \ F_c$$

Reject the null hypothesis. Therefore there the variances of NVDA and CMCSA stocks returns are significantly different from each other. T-test assuming unequal variances is appropriate in this case.

- **2. NVDA and EA**

$$H_0 : \sigma_N^2 = \sigma_E^2$$

$$H_1 : \sigma_N^2 \neq \sigma_E^2$$

$$\alpha = 0.01 \ \ then \ \frac{\alpha}{2} = 0.005$$

$$F = \left(\frac{\hat{\sigma}_N^2}{\hat{\sigma}_E^2}\right)$$

$$F = \frac{14.275}{6.752} = 2.114$$

$$F_c = 1.071$$

$$F > \ F_c$$

Hence reject the null hypothesis. Therefore NVDA and EA stocks variances are significantly different. We shall assume unequal variance when calculating t-value.

- **3. CMSA and EA**

$$H_0 : \sigma_C^2 = \sigma_E^2$$

$$H_1 : \sigma_C^2 \neq \sigma_E^2$$

$$\alpha = 0.01 \ \ then \ \ \frac{\alpha}{2} = 0.005$$

$$F = \left(\frac{\hat{\sigma}_C^2}{\hat{\sigma}_B^2}\right)$$

$$F = \frac{4.187}{6.752} = 0.062$$

$$F_c = 1.072$$

$$F < \ F_c$$

Fail to reject the null hypotheses This suggests that the variances returns of the two stocks returns does not differs significantly, hence t-test will assume equal variances.

$$Let\ \mu_N\ \ represent\ mean\ \ of\ \ NVDA\ \ stocks\ returns$$
$$let\ \mu_C\ \ =\ \ CMCSA\ average\ \ returns$$
$$let\ \mu_E\ =\ \ EA\ average\ returns$$

Perform t-test for the three pairs:

- **i) NVDA and CMCSA**

$$H_0 : \mu_N = \mu_C$$

$$H_1 : \mu_N \neq \mu_C$$

Consider that it's a two tailed t-test:

$$\alpha = 0.01\ then\ \frac{\alpha}{2} = 0.005$$

$$t = \frac{\hat{\mu}_N - \hat{\mu}_C}{\sqrt{\left(\frac{\hat{\sigma}_N^2}{N_N} + \frac{\hat{\sigma}_C^2}{N_C}\right)}}$$

computing using R:

```
NVDA_CMCS <- t.test(nvda_ret$per_returns, cmcsa_ret$per_returns, alternative = "two.sided",
    var.equal = F)
NVDA_CMCSA <- data.frame(NVDA_CMCS$statistic, NVDA_CMCS$p.value, NVDA_CMCS$parameter)
names(NVDA_CMCSA) <- c("t", "p-value", "df")
NVDA_CMCSA
```

```
##          t   p-value       df
## t 1.239161 0.2153192 8704.807
```

$$t = 1.2408$$
$$t_c = t_{df} = t_{8698.91}$$

Computing t critical using R:

```
qt(p = 0.005, df = 8698.91, lower.tail = FALSE)
```

```
## [1] 2.576395
```

$$t_c = 2.5764$$
$$|t|\ <\ |t_c|$$

In this case, fail to reject the null hypotheses and conclude that the average returns for NVDA and CMCSA are not significantly different.

- **ii) NVDA and EA**

$$H_0 : \mu_N = \mu_E$$

$$H_1 ; \mu_N \neq \mu_E$$

$$\alpha = 0.01 \; then \; \frac{\alpha}{2} = 0.005$$

$$t = \frac{\hat{\mu}_N - \hat{\mu}_E}{\sqrt{\left(\frac{\hat{\sigma}_N^2}{N_N} + \frac{\hat{\sigma}_E^2}{N_E}\right)}}$$

```
NVDA_E <- t.test(nvda_ret$per_returns, ea_ret$per_returns, alternative = "two.sided", var.equal = F)
NVDA_EA <- data.frame(NVDA_E$statistic, NVDA_E$p.value, NVDA_E$parameter)
names(NVDA_EA) <- c("t", "p-value", "df")
NVDA_EA
```

```
##         t   p-value       df
## t 1.02762 0.3041537 10016.41
```

$$t = 1.0378$$

$$t_c = t_{df} = t_{7455.11}$$

```
qt(p = 0.005, df = 10014.05, lower.tail = FALSE)
```

```
## [1] 2.57632
```

$$t_c = 2.5763$$

$$|t| \; < \; |t_c|$$

Fail to reject the null hypothesis and conclude that there is no significant difference between the average returns of NVDA and EA stocks.

- **iii) CMCSA and EA**

$$H_0 : \mu_H = \mu_N$$

$$H_1 ; \mu_H \neq \mu_N$$

$$\alpha = 0.01 \; then \; \frac{\alpha}{2} = 0.005$$

$$t = \frac{\hat{\mu}_C - \hat{\mu}_E}{\sqrt{\left(\frac{\hat{\sigma}_C^2}{N_C} + \frac{\hat{\sigma}_E^2}{N_E}\right)}}$$

Computing using R:

```
CMCSA_E <- t.test(cmcsa_ret$per_returns, ea_ret$per_returns, alternative = "two.sided", var.equal = T,
    )
CMCSA_EA <- data.frame(CMCSA_E$statistic, CMCSA_E$p.value, CMCSA_E$parameter)
names(CMCSA_EA) <- c("t", "p-value", "df")
CMCSA_EA
```

```
##           t   p-value    df
## t -0.185396 0.8529217 11300
```

$$t = -0.1732$$

$$t_c = t_{df} = t_{11296}$$

```
qt(p = 0.005, df = 11296, lower.tail = T)
```

```
## [1] -2.576265
```

$$t_c = -2.5763$$

$$t \ > \ t_c$$

Reject the null hypotheses and conclude that the average returns for CMCSA and EA stocks return are significantly different.

From the three tests of significance, the average stocks returns for CMCSA and EA are significantly different but both average returns for EA and CMSA are not significantly different from NVDA

2.7

```
corr <- data.frame(NVDA = nvda_ret$per_returns, CMCSA = cmcsa_ret$per_returns, EA = ea_ret$per_returns)
kable(cor(corr, method = "pearson"), caption = "Correlation matrix of the three returns")
```

Table 4: Correlation matrix of the three returns

	NVDA	CMCSA	EA
NVDA	1.0000000	0.3256291	0.3754527
CMCSA	0.3256291	1.0000000	0.3030105
EA	0.3754527	0.3030105	1.0000000

According to the correlation matrix, the three stocks returns are not strongly correlated with each other. The correlation coefficients between NVDA and CMCSA and EA are 0.3251 and 0.3754 respectively, while CMSA and EA pair has a coefficient of 0.3030.

2.8. Testing the significance of correlations

Set up the null and alternative hypothesis:

$$H_0 : \rho = 0$$

$$H_1 : \rho \neq 0$$

Consider the the following test statistic:

$$t = \frac{r\sqrt{N-2}}{\sqrt{1-r^2}}$$

Decision rule: If $p - value > 0$ reject the H_0

```
cor1 <- cor.test(nvda_ret$per_returns, cmcsa_ret$per_returns, method = "pearson")  #test between NVDA a
cor2 <- cor.test(nvda_ret$per_returns, ea_ret$per_returns, method = "pearson")  #test between NVDA and
cor3 <- cor.test(cmcsa_ret$per_returns, ea_ret$per_returns, method = "pearson")  # test between CMCSA a
```

```r
# add the outputs into a table
cor1 <- data.frame(statistic = cor1$statistic, p_value = round(cor1$p.value, 2))
row.names(cor1) <- c("NVDA_CMCSA")
cor2 <- data.frame(statistic = cor2$statistic, p_value = round(cor2$p.value, 2))
row.names(cor2) <- c("NVDA_EA")
cor3 <- data.frame(statistic = cor3$statistic, p_value = round(cor3$p.value, 2))
row.names(cor3) <- c("CMCSA_EA")
cor <- rbind(cor1, cor2, cor3)
kable(cor, caption = "Test for correlations significance")
```

Table 5: Test for correlations significance

	statistic	p_value
NVDA_CMCSA	25.88503	0
NVDA_EA	30.44637	0
CMCSA_EA	23.89772	0

Decision: Therefore p-values obtained equals to 0 which is less than $\alpha = 0.05$ significance level therefore reject the null hypotheses.

This indicates that the correlation between the stocks are significant.

2.9 Advising an investor.

Suppose that an investor has asked you to assist them in choosing **two** of these three stocks to include in their portfolio. The portfolio is defined by

$$r = w_1 r_1 + w_2 r_2$$

Where r_1 and r_2 represent the returns from the first and second stock, respectively, and w_1 and w_2 represent the proportion of the investment placed in each stock. The entire investment is allocated between the two stocks, so $w_1 + w_2 = 1$.

The investor favors the combination of stocks that provides the highest return, but dislikes risk. Thus the investor's happiness is a function of the portfolio, r:

$$h(r) = \mathbb{E}(r) - \mathbb{V}ar(r)$$

Where $\mathbb{E}(r)$is the expected return of the portfolio, and $\mathbb{V}ar(r)$ is the variance of the portfolio.[1]

FIRST of all, find the covariance matrix of returns

```r
Nvda <- NVDA_RT$per_returns
Cmcsa <- CMCSA_RT$per_returns
Ea <- EA_RT$per_returns
pfol <- data.frame(Nvda, Cmcsa, Ea)
kable(cov(pfol), caption = "covariance matrix")
```

[1]Note that $\mathbb{E}(r) = w_1 E(r_1) + w_2 \mathbb{E}(r_2)$, and $\mathbb{V}ar(r) = w_1^2 \mathbb{V}ar(r_1) + w_2^2 \mathbb{V}ar(r_2) + 2w_1 w_2 \mathbb{C}ov(r_1, r_2)$

Table 6: covariance matrix

	Nvda	Cmcsa	Ea
Nvda	14.276245	2.519131	3.685510
Cmcsa	2.519131	4.192189	1.611809
Ea	3.685510	1.611809	6.749487

Then calculate the optimum weights as follows:

$$h(r) = \mathbb{E}(r) - \mathbb{V}ar(r)$$

$$E(r) = E(w_1 r_1 + w_2 r_2) = w_1 E(r_1) + w_2 E(r_2)$$

$$Var(r) = Var(w_1 r_1 + w_2 r_2) = w_1 Var(r_1) + w_2 Var(r_2) + 2 w_1 w_2 COV(r_1, r_2)$$

$$Test\ For\ NVDA\ and\ CMCSA$$

$$E(r) = (0.0913)w_1 + (0.0204)w_2$$

$$Var(r) = 14.2747w_1 + 4.1922w_2 + 2(2.519)w_1 w_2$$

$$w_1 + w_2 = 1$$

$$w_2 = 1 - w_1$$

$$E(r) = 0.0913w_1 + 0.0204(1 - w_1)$$

$$= 0.071w_1 + 0.0204$$
$$Var(r) = 14.2747w_1 + 4.1922(1 - w_1) + 2(2.519)w_1(1 - w_1)$$
$$= 14.2747w_1 + 4.1922 - 4.1922w_1 + 5.038w_1 - 5.038w_1^2$$
$$= 4.1922 + 15.1207w_1 - 5.038w_1^2$$
$$h(r) = 0.071w_1 + 0.0204 - (4.1922 + 15.1207w_1 - 5.038w_1^2)$$
$$= 0.01w_1 + 0.047 - 1.7 - 3.92w_1 + 1.82w_1^2$$
$$= 5.038w_1^2 - 15.0497w_1 - 4.1718$$
$$first\ order\ derivative\ = \frac{dh(r)}{dw_1} = 10.076w_1 - 15.0497$$
$$for\ optimum\ weight,\ 10.076w_1 - 14.0436 = 0$$
$$therefore,\ w_1 = 1.494,\ and\ w_2 = 1 - w_1 = -0.494$$
$$second\ order\ derivative\ = \frac{dh(r)^2}{d^2 w_1} = 10.054$$

$$\frac{dh(r)^2}{d^2w_1} > 0, opt\ weight\ is\ minimum.$$

$$Hence\ the\ expected\ return\ for\ portfolio\ NVDA\ and\ CMCSA\ is$$

$$E(r) = (0.0913)(1.494) + (0.0204)(-0.494) = 0.1263$$

$$Var(r) = 14.2747(1.4044) + 4.1871(-0.4044) + 2(1.4044)(-0.4044)(2.5135) = 7.1795$$

$$h(r) = 0.1211 - 7.1795 = -7.0584$$

$$TEST\ FOR\ NVDA\ and\ EA$$

$$E(r) = (0.0924)w_1 + (0.0291)w_2$$

$$E(r) = (0.0924)w_1 + (0.0291) - 0.0291w_1$$

$$Var(r) = 14.2747w_1 + 6.7517w_2 + 2(3.6856)w_1w_2$$

$$w_1 + w_2 = 1$$

$$w_2 = 1 - w_1$$

$$E(r) = 0.0924w_1 + 0.0291(1 - w_1)$$

$$= 0.0924w_1 + 0.0291 - 0.0291w_1$$

$$= 0.0633w_1 + 0.0291$$

$$Var(r) = 14.2747w_1 + 6.7517(1 - w_1) + 2w_1(3.6856)(1 - w_1)$$

$$= 14.2747w_1 + 6.7517 - 6.7517w_1 + 7.3712w_1 - 7.3712w_1^2)$$

$$= 7.3712w_1^2 + 14.8942w_1 - 6.7517$$

$$h(r) = 0.0633w_1 + 0.0291 - (7.3712w_1^2 + 14.8942w_1 - 6.7517)$$

$$= 6.7808 - 14.8309w_1 - 7.3712w_1^2$$

$$first\ order\ derivative\ = \frac{dh(r)}{dw_1} = -14.7424w_1 - 14.8309$$

$$for\ optimum\ weight,\ -14.7424w_1 - 14.8309 = 0$$

$$therefore,\ w_1 = -1.006,\ and\ w_2 = 1 - w_1 = 2.006$$

$$second\ order\ derivative\ = \frac{dh(r)^2}{d^2w_1} = -14.7424$$

$$\frac{dh(r)^2}{d^2w_1} < 0, opt\ weight\ is\ maximum.$$

$$Hence\ the\ expected\ return\ for\ portfolio\ NVDA\ and\ EA\ is$$

$$SUBSTITUTE\ w_1\ and\ w_2\ to\ E(r),\ Var(r),\ and\ h(r)$$

$$E(r) = (0.0924)(-1.006) + (0.0291)(2.006) = -0.0346$$

$$Var(r) = 14.2747(-1.006) + 6.7517(2.006) + 2(3.6856)(-1.006)(2.006) = -15.69$$

$$h(r) = -0.0346 + 15.69 = 15.66$$

$$TEST\ FOR\ CMCSA\ and\ EA$$

$$E(r) = (0.0214)w_1 + (0.0291)w_2$$

$$Var(r) = 4.1871w_1 + 6.7517w_2 + 2(1.6111)w_1w_2$$

$$w_1 + w_2 = 1$$

$$w_2 = 1 - w_1$$

$$E(r) = (0.0214)w_1 + (0.0291)(1 - w_1)$$

$$= 0.0291 - 0.0077w_1$$

$$Var(r) = 4.1871w_1 + 6.7517(1 - w_1) + 2w_1(1.6111)(1 - w_1)$$

$$= 0.6576 + 6.7517w_1 - 3.222w_1^2$$

$$h(r) = 0.0291 - 0.0077w_1 - (0.6576 + 6.7517w_1 - 3.222w_1^2)$$

$$= 3.222w_1^2 - 6.744w_1 - 0.6285$$

$$first\ order\ derivative\ = \frac{dh(r)}{dw_1} = 6.444w_1 - 6.744$$

$$for\ optimum\ weight,\ 6.444w_1 - 6.744 = 0$$

$$therefore,\ w_1 = 1.047,\ and\ w_2 = 1 - w_1 = -0.0466$$

$$second\ order\ derivative\ = \frac{dh(r)^2}{d^2w_1} = 6.444$$

$$\frac{dh(r)^2}{d^2w_1} > 0, opt\ weight\ is\ minimum.$$

$$Therefore$$

$$E(r) = 0.0214(1.047) + (0.0291)(-0.0466) = 0.02105$$

$$Var(r) = 4.1871(1.047) + 6.7517(-0.0466) + 2(1.6111)(1.047)(-0.0466) = 3.912$$

$$h(r) = 0.02105 - 3.912 = -3.8909$$

```r
# You can use this section to create a table of your results.
optimal_pfolio <- matrix(c(1.494, -0.494, 0.1263, 7.1796, -7.0584, -1.006, 2.006, -0.0346,
    -15.69, 15.66, 1.047, -0.0466, 0.0211, 3.912, -3.912), ncol = 5, byrow = TRUE)
colnames(optimal_pfolio) <- c("weight_1", "weight_2", "E(r)", "Var(r)", "h(r)")
rownames(optimal_pfolio) <- c("NVDA and CMCSA", "NVDA and EA", "CMCSA and EA")
optimal_pfolio <- as.table(optimal_pfolio)
kable(optimal_pfolio, caption = "Optimum portfolio table")
```

Table 7: Optimum portfolio table

	weight_1	weight_2	E(r)	Var(r)	h(r)
NVDA and CMCSA	1.494	-0.4940	0.1263	7.1796	-7.0584
NVDA and EA	-1.006	2.0060	-0.0346	-15.6900	15.6600
CMCSA and EA	1.047	-0.0466	0.0211	3.9120	-3.9120

According the expected returns, A combination of NVDA and CMCSA generates the biggest returns with a 0.1263 coefficient. This implies that an investor has a higher probability(12.63%) of generating a profit. Since the investor is interested in a strategy with less uncertainty, CMCSA could provide a better option since it has the lowest variations in returns, hence more stable portfolio. However, the weights of the the two stocks (CMCSA and EA) suggests that the investor needs to consider investing 4.66% in EA and 95.34% in CMCSA.

2.10. The impact of financial events on returns (6 points)

```r
extract <- function(stocks, smb) {
    stoc <- subset(stocks, symbol == smb)
    lehman <- stoc[stoc$date == "2008-09-15", ]  # Lehman bankruptcy
    pandemic <- stoc[stoc$date == "2020-03-11", ]  # COVID pandemic
    bankrpcy1 <- stoc[stoc$date != "2008-09-15", ]  #omit the bankruptcy date
    bankrpcy2 <- stoc[stoc$date != "2020-03-11", ]  #omit the pandemic date
    BAU <- rbind(bankrpcy1, bankrpcy2)  #joining the dataframes
    # create new column with three factors
    BAU$event <- "BAU"
    lehman$event = "lehman"
    pandemic$event <- "pandemic"
    events <- rbind(lehman, pandemic, BAU)  #merge the three dataframes
    events$event <- factor(events$event)  #column to a factor
    events <- na.omit(events)  #drop the na columns
    return(events)
}

# call the function to extract and create the factor column
NVDA_EVENTS <- extract(stocks, "NVDA")
CMCSA_EVENTS <- extract(stocks, "CMCSA")
EA_EVENTS <- extract(stocks, "EA")
# linear regression model for the three stocks
model_nvda <- lm(per_returns ~ event, data = NVDA_EVENTS)
model_cmcsa <- lm(per_returns ~ event, dat = CMCSA_EVENTS)
model_ea <- lm(per_returns ~ event, data = EA_EVENTS)

# NVDA events
summary(model_nvda)$coefficients
```

```
##                  Estimate Std. Error   t value    Pr(>|t|)
## (Intercept)    0.09296391 0.03553253  2.616304 0.008900557
## eventlehman   -9.23219407 3.77732699 -2.444108 0.014536272
## eventpandemic -5.85163033 3.77732699 -1.549146 0.121374676
```

```r
# CMCSA events
summary(model_cmcsa)$coefficients
```

```
##                    Estimate Std. Error   t value   Pr(>|t|)
## (Intercept)    0.02141967 0.01925671  1.112322 0.26602326
## eventlehman   -4.01537863 2.04710692 -1.961489 0.04984647
## eventpandemic -3.11214962 2.04710692 -1.520267 0.12847180
```

```r
# EA events
summary(model_ea)$coefficients
```

```
##                    Estimate Std. Error   t value  Pr(>|t|)
## (Intercept)    0.02948126 0.02443765  1.206387 0.2276937
## eventlehman   -3.85823240 2.59787296 -1.485151 0.1375319
## eventpandemic -2.18051938 2.59787296 -0.839348 0.4012918
```

(a) According to the regression models, all the three stocks never exhibited positive returns over time since R-squared from the three models are not significantly different from zero.

(b) Both pandemic and bankruptcy events have negative impacts on the three stocks. In all cases, Lehman Brothers bankruptcy affected the stocks more negatively as the "eventlehman" coefficients are less than "eventpandemic" coefficients. Lehman bankruptcy had significantly negative impacts on both NVDA($p-value = 0.0146$) and CMCSA($p-value = 0.0499$) at 5% significance level.