Assignment 4

Bailey Scoville, Anhua Cheng, Griffin Arnone

2024-02-22

```
start_time <- Sys.time()

data_active_1 <- read.csv("PCCOX_R.csv", header = TRUE, sep = ",", stringsAsFactors = TRUE)
data_active_2 <- read.csv("PRILX.csv", header = TRUE, sep = ",", stringsAsFactors = TRUE)
data_active_3 <- read.csv("RWMGX.csv", header = TRUE, sep = ",", stringsAsFactors = TRUE)
data_passive <- read.csv("WFSPX.csv", header = TRUE, sep = ",", stringsAsFactors = TRUE)

#inspect data
head(data_active_1)
```

Import Fund Data

```
##       Date  Open  High   Low Close Adj.Close Volume
## 1 11/29/16 23.88 23.88 23.88 23.88  19.37983      0
## 2 11/30/16 23.77 23.77 23.77 23.77  19.29056      0
## 3  12/1/16 23.64 23.64 23.64 23.64  19.18506      0
## 4  12/2/16 23.66 23.66 23.66 23.66  19.20129      0
## 5  12/6/16 23.87 23.87 23.87 23.87  19.37172      0
## 6  12/7/16 24.17 24.17 24.17 24.17  19.61518      0
```

```
str(data_active_1)
```

```
## 'data.frame':    1815 obs. of  7 variables:
##  $ Date     : Factor w/ 1815 levels "1/10/17","1/10/18",..: 416 428 465 526 605 612 619 625 480 486
##  $ Open     : num  23.9 23.8 23.6 23.7 23.9 ...
##  $ High     : num  23.9 23.8 23.6 23.7 23.9 ...
##  $ Low      : num  23.9 23.8 23.6 23.7 23.9 ...
##  $ Close    : num  23.9 23.8 23.6 23.7 23.9 ...
##  $ Adj.Close: num  19.4 19.3 19.2 19.2 19.4 ...
##  $ Volume   : int  0 0 0 0 0 0 0 0 0 0 ...
```

```
head(data_active_2)
```

```
##        Date  Open  High   Low Close Adj.Close Volume
## 1 4/28/2006 25.59 25.59 25.59 25.59  9.497851      0
## 2  5/1/2006 25.51 25.51 25.51 25.51  9.468159      0
## 3  5/2/2006 25.58 25.58 25.58 25.58  9.494144      0
## 4  5/3/2006 25.61 25.61 25.61 25.61  9.505277      0
## 5  5/4/2006 25.67 25.67 25.67 25.67  9.527547      0
## 6  5/5/2006 25.85 25.85 25.85 25.85  9.594355      0
```

```
str(data_active_2)
```

```
## 'data.frame':    4482 obs. of  7 variables:
##  $ Date     : Factor w/ 4482 levels "1/10/2007","1/10/2008",..: 2453 2571 2714 2841 2875 2889 2926 2
```

```
##  $ Open     : num  25.6 25.5 25.6 25.6 25.7 ...
##  $ High     : num  25.6 25.5 25.6 25.6 25.7 ...
##  $ Low      : num  25.6 25.5 25.6 25.6 25.7 ...
##  $ Close    : num  25.6 25.5 25.6 25.6 25.7 ...
##  $ Adj.Close: num  9.5 9.47 9.49 9.51 9.53 ...
##  $ Volume   : int  0 0 0 0 0 0 0 0 0 0 ...
```

```
head(data_active_3)
```

```
##       Date  Open  High   Low Close Adj.Close Volume
## 1 5/1/2009 19.95 19.95 19.95 19.95  8.865779      0
## 2 5/4/2009 20.58 20.58 20.58 20.58  9.145751      0
## 3 5/5/2009 20.52 20.52 20.52 20.52  9.119089      0
## 4 5/6/2009 20.79 20.79 20.79 20.79  9.239075      0
## 5 5/7/2009 20.55 20.55 20.55 20.55  9.132417      0
## 6 5/8/2009 21.00 21.00 21.00 21.00  9.332400      0
```

```
str(data_active_3)
```

```
## 'data.frame':    3725 obs. of  7 variables:
##  $ Date     : Factor w/ 3725 levels "1/10/2011","1/10/2012",..: 2133 2385 2397 2408 2419 2429 2154 2
##  $ Open     : num  20 20.6 20.5 20.8 20.5 ...
##  $ High     : num  20 20.6 20.5 20.8 20.5 ...
##  $ Low      : num  20 20.6 20.5 20.8 20.5 ...
##  $ Close    : num  20 20.6 20.5 20.8 20.5 ...
##  $ Adj.Close: num  8.87 9.15 9.12 9.24 9.13 ...
##  $ Volume   : int  0 0 0 0 0 0 0 0 0 0 ...
```

```
head(data_passive)
```

```
##        Date  Open  High   Low Close Adj.Close Volume
## 1  7/2/1993 80.00 80.00 80.00 80.00  19.30723      0
## 2  7/6/1993 79.20 79.20 79.20 79.20  19.11415      0
## 3  7/7/1993 79.44 79.44 79.44 79.44  19.17208      0
## 4  7/8/1993 80.48 80.48 80.48 80.48  19.42306      0
## 5  7/9/1993 80.40 80.40 80.40 80.40  19.40377      0
## 6 7/12/1993 80.56 80.56 80.56 80.56  19.44238      0
```

```
str(data_passive)
```

```
## 'data.frame':    7712 obs. of  7 variables:
##  $ Date     : Factor w/ 7712 levels "1/10/1994","1/10/1995",..: 5985 6306 6329 6352 6374 5807 5830 58
##  $ Open     : num  80 79.2 79.4 80.5 80.4 ...
##  $ High     : num  80 79.2 79.4 80.5 80.4 ...
##  $ Low      : num  80 79.2 79.4 80.5 80.4 ...
##  $ Close    : num  80 79.2 79.4 80.5 80.4 ...
##  $ Adj.Close: num  19.3 19.1 19.2 19.4 19.4 ...
##  $ Volume   : int  0 0 0 0 0 0 0 0 0 0 ...
```

```
# Record end time
end_time <- Sys.time()
execution_time <- end_time - start_time
cat(paste("Execution Time: ", execution_time, " seconds\n"))
```

```
## Execution Time:  0.0602250099182129  seconds
```

```
# Get memory usage
memory_info <- pryr::mem_used()
```

```r
cat(paste("Memory Usage: ", memory_info / 1024 / 1024, " MB\n"))
```

```
## Memory Usage:  59.3085861206055  MB
```

```r

start_time <- Sys.time()

pct_change1 <- diff(data_active_1$Adj.Close) / lag(data_active_1$Adj.Close)
adj_change1 <- 1 + pct_change1
log_return_active_1 <- log(adj_change1)

pct_change2 <- diff(data_active_2$Adj.Close) / lag(data_active_2$Adj.Close)
adj_change2 <- 1 + pct_change2
log_return_active_2 <- log(adj_change2)

pct_change3 <- diff(data_active_3$Adj.Close) / lag(data_active_3$Adj.Close)
adj_change3 <- 1 + pct_change3
log_return_active_3 <- log(adj_change3)

pct_change_passive <- diff(data_passive$Adj.Close) / lag(data_passive$Adj.Close)
adj_change_passive <- 1 + pct_change_passive
log_return_passive <- log(adj_change_passive)

# Record end time
end_time <- Sys.time()
execution_time <- end_time - start_time
cat(paste("Execution Time: ", execution_time, " seconds\n"))
```

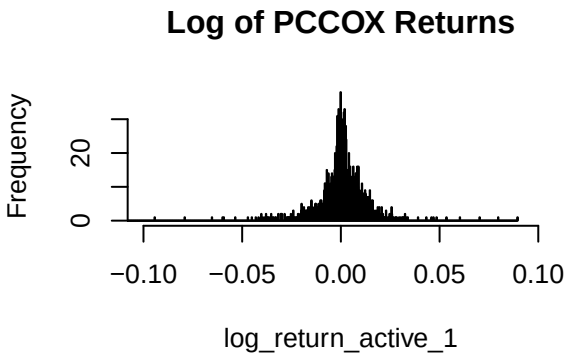
Data Transformation

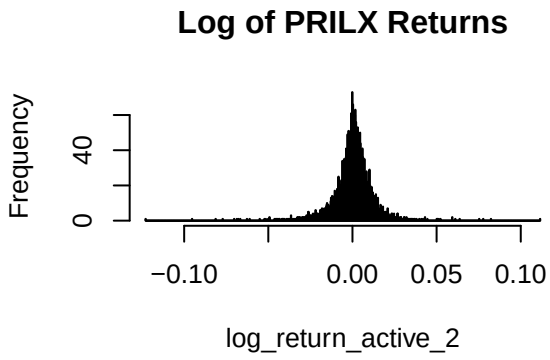
```
## Execution Time:  0.0040888786315918  seconds
```

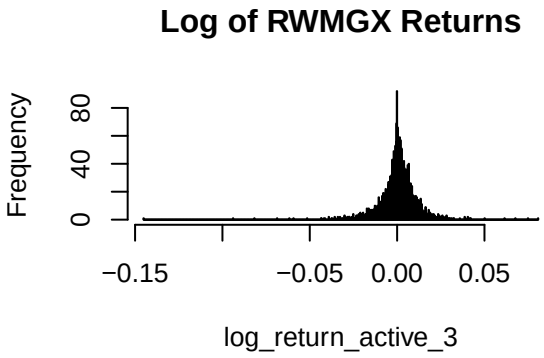
```r
# Get memory usage
memory_info <- pryr::mem_used()
cat(paste("Memory Usage: ", memory_info / 1024 / 1024, " MB\n"))
```

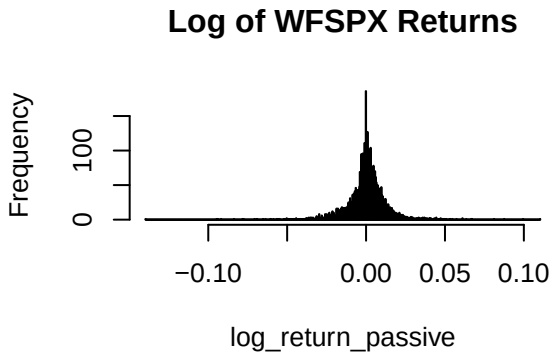
```
## Memory Usage:  59.8068008422852  MB
```

```r
start_time <- Sys.time()

par(mfrow = c(2,2))

hist(log_return_active_1, breaks = 1000, main = "Log of PCCOX Returns", xlim = c(-0.1, 0.1))
hist(log_return_active_2, breaks = 1000,  main = "Log of PRILX Returns")
hist(log_return_active_3, breaks = 1000,  main = "Log of RWMGX Returns")
hist(log_return_passive, breaks = 1000,  main = "Log of WFSPX Returns")
```










```r
par(mfrow = c(1,1))

# Record end time
end_time <- Sys.time()
execution_time <- end_time - start_time
cat(paste("Execution Time: ", execution_time, " seconds\n"))
```

```
## Execution Time:  0.0055391788482666  seconds
```

```r
# Get memory usage
memory_info <- pryr::mem_used()
cat(paste("Memory Usage: ", memory_info / 1024 / 1024, " MB\n"))
```

```
## Memory Usage:  60.4217758178711  MB
```

```r

start_time <- Sys.time()

mean_active_1 <- mean(log_return_active_1, na.rm = T)
var_active_1 <- var(log_return_active_1, na.rm = T)
drift_active_1 <- mean_active_1 - (0.5 * var_active_1)

mean_active_2 <- mean(log_return_active_2, na.rm = T)
var_active_2 <- var(log_return_active_2, na.rm = T)
drift_active_2 <- mean_active_2 - (0.5 * var_active_2)

mean_active_3 <- mean(log_return_active_3, na.rm = T)
var_active_3 <- var(log_return_active_3, na.rm = T)
drift_active_3 <- mean_active_3 - (0.5 * var_active_3)
```

```r
mean_passive <- mean(log_return_passive, na.rm = T)
var_passive <- var(log_return_passive, na.rm = T)
drift_passive <- mean_passive - (0.5 * var_passive)

# Record end time
end_time <- Sys.time()
execution_time <- end_time - start_time
cat(paste("Execution Time: ", execution_time, " seconds\n"))
```

Simulation

```
## Execution Time:  0.00346016883850098  seconds
```

```r
# Get memory usage
memory_info <- pryr::mem_used()
cat(paste("Memory Usage: ", memory_info / 1024 / 1024, " MB\n"))
```

```
## Memory Usage:  60.4385375976562  MB
```

```r
start_time <- Sys.time()

days <- 251
trials <- 10000

stdev_active_1 <- sd(log_return_active_1, na.rm = T)
Z_active_1 <- matrix(qnorm(runif(days * trials)), nrow = days, ncol = trials)
daily_returns_active_1 <- exp(drift_active_1 + stdev_active_1 * Z_active_1)

stdev_active_2 <- sd(log_return_active_2, na.rm = T)
Z_active_2 <- matrix(qnorm(runif(days * trials)), nrow = days, ncol = trials)
daily_returns_active_2 <- exp(drift_active_2 + stdev_active_2 * Z_active_2)

stdev_active_3 <- sd(log_return_active_3, na.rm = T)
Z_active_3 <- matrix(qnorm(runif(days * trials)), nrow = days, ncol = trials)
daily_returns_active_3 <- exp(drift_active_3 + stdev_active_3 * Z_active_3)

stdev_passive <- sd(log_return_passive, na.rm = T)
Z_passive  <- matrix(qnorm(runif(days * trials)), nrow = days, ncol = trials)
daily_returns_passive <- exp(drift_passive  + stdev_passive  * Z_passive)

# Record end time
end_time <- Sys.time()
execution_time <- end_time - start_time
cat(paste("Execution Time: ", execution_time, " seconds\n"))
```

```
## Execution Time:  0.388107061386108  seconds
```

```r
# Get memory usage
memory_info <- pryr::mem_used()
cat(paste("Memory Usage: ", memory_info / 1024 / 1024, " MB\n"))
```

```
## Memory Usage:  213.655944824219  MB
```

```r
start_time <- Sys.time()

price_paths_active_1 <- matrix(0, nrow = days, ncol = trials)
price_paths_active_1[1, ] <- data_active_1[nrow(data_active_1), 6]
```

```r
for (t in 2:days) {
  price_paths_active_1[t, ] <- price_paths_active_1[t-1, ] * (daily_returns_active_1[t, ])
}

price_paths_active_2 <- matrix(0, nrow = days, ncol = trials)
price_paths_active_2[1, ] <- data_active_2[nrow(data_active_2), 6]
for (t in 2:days) {
  price_paths_active_2[t, ] <- price_paths_active_2[t-1, ] * (daily_returns_active_2[t, ])
}

price_paths_active_3 <- matrix(0, nrow = days, ncol = trials)
price_paths_active_3[1, ] <- data_active_3[nrow(data_active_3), 6]
for (t in 2:days) {
  price_paths_active_3[t, ] <- price_paths_active_3[t-1, ] * (daily_returns_active_3[t, ])
}

price_paths_passive <- matrix(0, nrow = days, ncol = trials)
price_paths_passive[1, ] <- data_passive[nrow(data_passive), 6]
for (t in 2:days) {
  price_paths_passive[t, ] <- price_paths_passive[t-1, ] * (daily_returns_passive[t, ])
}

# Record end time
end_time <- Sys.time()
execution_time <- end_time - start_time
cat(paste("Execution Time: ", execution_time, " seconds\n"))
```

```
## Execution Time:  0.297363996505737  seconds
```

```r
# Get memory usage
memory_info <- pryr::mem_used()
cat(paste("Memory Usage: ", memory_info / 1024 / 1024, " MB\n"))
```

```
## Memory Usage:  290.331611633301  MB
```

```r
#inspect price path array
dim(price_paths_active_1)
```

```
## [1]   251 10000
```

```r
start_time <- Sys.time()

ncol <- 251
col_names <- paste('Day', 1:ncol, sep = "_")

#price path 1
df1 <- as.data.frame(price_paths_active_1)
df1 <- t(df1)
colnames(df1) <- col_names

#price path 2
df2 <- as.data.frame(price_paths_active_2)
df2 <- t(df2)
colnames(df2) <- col_names

#price path 3
```

```r
df3 <- as.data.frame(price_paths_active_3)
df3 <- t(df3)
colnames(df3) <- col_names

#price path passive
df_passive <- as.data.frame(price_paths_passive)
df_passive <- t(df_passive)
colnames(df_passive) <- col_names

# Record end time
end_time <- Sys.time()
execution_time <- end_time - start_time
cat(paste("Execution Time: ", execution_time, " seconds\n"))
```

```
## Execution Time:  0.225641012191772  seconds
```

```r
# Get memory usage
memory_info <- pryr::mem_used()
cat(paste("Memory Usage: ", memory_info / 1024 / 1024, " MB\n"))
```

```
## Memory Usage:  367.897407531738  MB
```

```r
start_time <- Sys.time()

#PCCOX - data active 1
df1 <- transform(df1, Volatility=apply(df1, 1, sd, na.rm=TRUE))
df1$Return <- (df1$Day_251 - df1$Day_1) / df1$Day_1

#PRILX - data active 2
df2 <- transform(df2, Volatility=apply(df2, 1, sd, na.rm=TRUE))
df2$Return <- (df2$Day_251 - df2$Day_1) / df2$Day_1

#RWMGX - data active 3
df3 <- transform(df3, Volatility=apply(df3, 1, sd, na.rm=TRUE))
df3$Return <- (df3$Day_251 - df3$Day_1) / df3$Day_1

#WFSPX - data passive
df_passive <- transform(df_passive, Volatility=apply(df_passive, 1, sd, na.rm=TRUE))
df_passive$Return <- (df_passive$Day_251 - df_passive$Day_1) / df_passive$Day_1

# Record end time
end_time <- Sys.time()
execution_time <- end_time - start_time
cat(paste("Execution Time: ", execution_time, " seconds\n"))
```

```
## Execution Time:  0.515664100646973  seconds
```

```r
# Get memory usage
memory_info <- pryr::mem_used()
cat(paste("Memory Usage: ", memory_info / 1024 / 1024, " MB\n"))
```

```
## Memory Usage:  368.63321685791  MB
```

```r
start_time <- Sys.time()

#average returns
```

```r
PCCOX_returns <- mean(df1$Return)
PRILX_returns <- mean(df2$Return)
RWMGX_returns <- mean(df3$Return)
WFSPX_returns <- mean(df_passive$Return)

#create vectors
Avg_Annual_Return <- c(PCCOX_returns, PRILX_returns, RWMGX_returns, WFSPX_returns)
Fund <- c("PCCOX", "PRILX", "RWMGX", "WFSPX")

#create returns table
returns_tab <- cbind.data.frame(Fund, Avg_Annual_Return)
returns_tab
```

```
##    Fund Avg_Annual_Return
## 1 PCCOX         0.1335404
## 2 PRILX         0.1011866
## 3 RWMGX         0.1349377
## 4 WFSPX         0.1162798
```

```r
# Record end time
end_time <- Sys.time()
execution_time <- end_time - start_time
cat(paste("Execution Time: ", execution_time, " seconds\n"))
```

```
## Execution Time:  0.00264811515808105  seconds
```

```r
# Get memory usage
memory_info <- pryr::mem_used()
cat(paste("Memory Usage: ", memory_info / 1024 / 1024, " MB\n"))
```

```
## Memory Usage:  368.628311157227  MB
```

```r
start_time <- Sys.time()

write.csv(df1, 'PCCOX_returns_R.csv', row.names = FALSE)
write.csv(df2, 'PRILX__returns_R.csv', row.names = FALSE)
write.csv(df3, 'RWMGX_returns_R.csv', row.names = FALSE)
write.csv(df_passive, 'WFSPX_returns_R.csv', row.names = FALSE)

# Record end time
end_time <- Sys.time()
execution_time <- end_time - start_time
cat(paste("Execution Time: ", execution_time, " seconds\n"))
```

```
## Execution Time:  4.92530703544617  seconds
```

```r
# Get memory usage
memory_info <- pryr::mem_used()
cat(paste("Memory Usage: ", memory_info / 1024 / 1024, " MB\n"))
```

```
## Memory Usage:  368.73063659668  MB
```