README

20903936

12/4/2021

Question 3:

In this brief report, I aim to show the JSE the concentration and commonality of returns within the Top 40 index (J200). In order to do so, I will look at using Principal Component Analysis and rolling constituent correlation perspective to illustrate. In my answer I will use stratification of monthly volatility by calculating the J200 index returns and comparing return source concentrations for periods of high volatility only.

```
# Importing the data
T40 <-readRDS("/Users/mathlogonolomashitisho/Desktop/Economics Masters /Semester 2/Fin Metrics/Financial
```

Let begin by stratifying the returns for times during high volatility.

```
library(rmsfuns)
pacman::p_load("tidyr", "tbl2xts","devtools","lubridate", "readr", "PerformanceAnalytics", "ggplot2", "d
```

```
#Here , there is no need to consider before 2008 because the data has already provided for the survivin

T40 <- T40 %>% group_by(date) %>% dplyr::mutate(Return= coalesce(Return,0)) %>% dplyr::mutate(J200= coal
```

```
T40 <-

  T40 %>% group_by(Tickers) %>%

  mutate(Top = quantile(J200return, 0.99), Bot = quantile(J200return, 0.01)) %>%

  mutate(Return = ifelse(J200return> Top, Top,

                         ifelse(J200return < Bot, Bot, J200return))) %>% ungroup() %>%  mutate(YearMont
```

```
T40SD <-T40  %>%

  mutate(YearMonth = format(date, "%Y%B")) %>%

  group_by(YearMonth) %>% summarise(SD = sd(J200return)*sqrt(52)) %>%

  # Top Decile Quantile overall (highly volatile month for ZAR:
  mutate(TopQtile = quantile(SD, 0.8),

         BotQtile = quantile(SD, 0.2))
```

```
Hi_Vol <- T40SD %>% filter(SD > TopQtile) %>% pull(YearMonth)

Low_Vol <- T40SD %>% filter(SD < BotQtile) %>% pull(YearMonth)
```

```
# Create generic function to compare performance:

Perf_comparisons <- function(Idxs, YMs, Alias){
  # For stepping through uncomment:
  YMs <- Hi_Vol

  Unconditional_SD <-

  Idxs %>%

    group_by(Tickers) %>%

    mutate(Full_SD = sd(Return) * sqrt(252)) %>%

    filter(YearMonth %in% YMs) %>%

    summarise(SD = sd(Return) * sqrt(252), across(.cols = starts_with("Full"), .fns = max)) %>%

    arrange(desc(SD)) %>% mutate(Period = Alias) %>%

    group_by(Tickers) %>%

    mutate(Ratio = SD / Full_SD)

    Unconditional_SD

}
```

```
#Extract the Tickers with monthly high volatility

perf_hi <- Perf_comparisons(T40, YMs = Hi_Vol, Alias = "High_Vol")

perf_lo <- Perf_comparisons(T40, YMs = Low_Vol, Alias = "Low_Vol")
```

PCA Analysis:

```
#Removing rows with missing values

perf_hi <-na.omit(perf_hi)
```

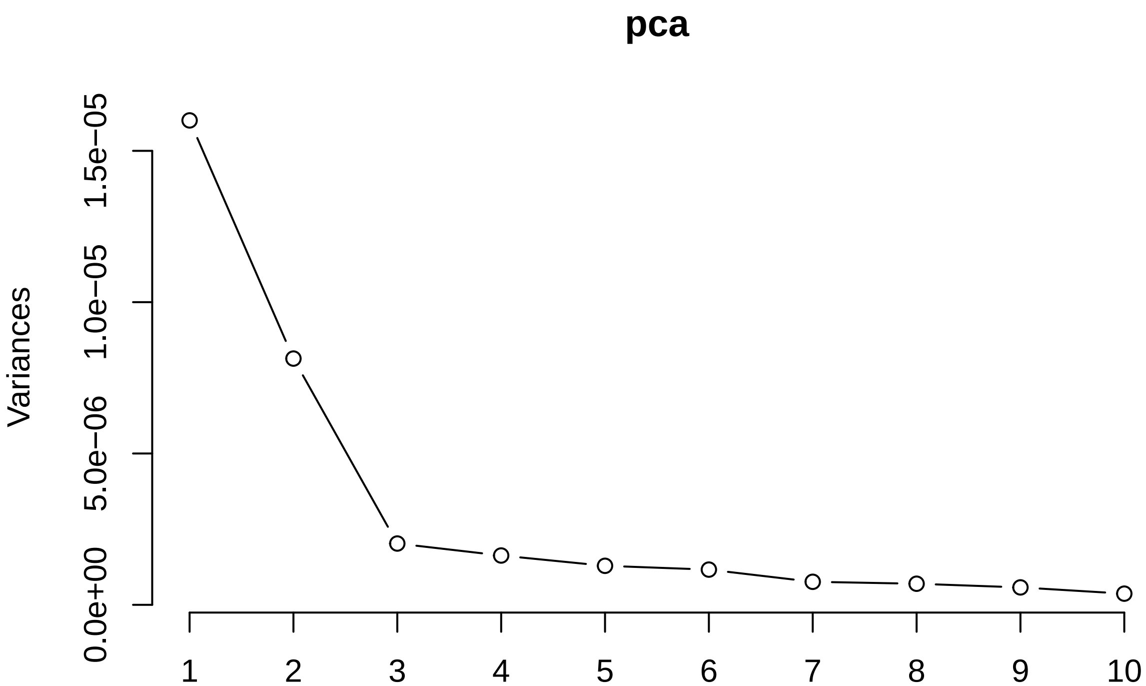
Mean centering J200returns

```
T40_Centered <-T40 %>% group_by(Tickers) %>% mutate(J200return_centered = J200return  -
    mean(J200return)) %>% ungroup()
```

Here we are only interested in Hi volatility periods as stratified above.

```r
pacman::p_load(fEcofin)

# prcomp requires wide, numeric data:

data_wide <- T40_Centered %>% filter(T40_Centered$Tickers %in% perf_hi$Tickers) %>% dplyr::mutate(Return
    J200return_centered) %>% select(-date)


# make sure to account for only those in portfolio
temp<-data_wide[, colSums(is.na(data_wide)) != nrow(data_wide)]

data_wide <- temp %>%
    mutate(
        across(everything(), replace_na, 0)
    )

pca <- prcomp(data_wide)
# We have already cebtered our data, but we could also use
# these built in measures: prcomp(data_wide, center = TRUE,
# scale. = TRUE)
```

```r
plot(pca, type = "l")
```




```r
summary(pca)
```

```
## Importance of components:
##                             PC1      PC2      PC3      PC4      PC5      PC6
## Standard deviation     0.004001 0.002852 0.001423 0.001276 0.001135 0.001079
## Proportion of Variance 0.467120 0.237460 0.059110 0.047560 0.037620 0.033960
## Cumulative Proportion  0.467120 0.704580 0.763700 0.811250 0.848870 0.882840
```

```
##                              PC7       PC8       PC9      PC10      PC11
## Standard deviation     0.0008722 0.0008349 0.0007587 0.0006079 0.0004087
## Proportion of Variance 0.0222000 0.0203400 0.0168000 0.0107900 0.0048700
## Cumulative Proportion  0.9050400 0.9253800 0.9421800 0.9529600 0.9578400
##                             PC12      PC13      PC14      PC15     PC16
## Standard deviation     0.0003829 0.0003571 0.0003364 0.0003143 0.000282
## Proportion of Variance 0.0042800 0.0037200 0.0033000 0.0028800 0.002320
## Cumulative Proportion  0.9621100 0.9658400 0.9691400 0.9720200 0.974340
##                             PC17      PC18      PC19     PC20      PC21
## Standard deviation     0.0002768 0.0002703 0.0002572 0.000231 0.0002264
## Proportion of Variance 0.0022400 0.0021300 0.0019300 0.001560 0.0015000
## Cumulative Proportion  0.9765800 0.9787100 0.9806400 0.982200 0.9836900
##                             PC22      PC23      PC24      PC25      PC26
## Standard deviation     0.0002146 0.0002088 0.0001835 0.0001755 0.0001707
## Proportion of Variance 0.0013400 0.0012700 0.0009800 0.0009000 0.0008500
## Cumulative Proportion  0.9850400 0.9863100 0.9872900 0.9881900 0.9890400
##                             PC27      PC28      PC29      PC30      PC31
## Standard deviation     0.0001667 0.0001534 0.0001494 0.0001406 0.0001396
## Proportion of Variance 0.0008100 0.0006900 0.0006500 0.0005800 0.0005700
## Cumulative Proportion  0.9898500 0.9905400 0.9911900 0.9917700 0.9923400
##                             PC32      PC33      PC34     PC35     PC36
## Standard deviation     0.0001355 0.0001313 0.0001274 0.000125 0.000119
## Proportion of Variance 0.0005400 0.0005000 0.0004700 0.000460 0.000410
## Cumulative Proportion  0.9928700 0.9933800 0.9938500 0.994310 0.994720
##                             PC37      PC38      PC39      PC40      PC41
## Standard deviation     0.0001164 0.0001123 0.0001066 0.0001058 0.0001013
## Proportion of Variance 0.0004000 0.0003700 0.0003300 0.0003300 0.0003000
## Cumulative Proportion  0.9951200 0.9954800 0.9958200 0.9961400 0.9964400
##                             PC42      PC43      PC44      PC45      PC46
## Standard deviation     9.439e-05 9.071e-05 8.701e-05 8.538e-05 8.122e-05
## Proportion of Variance 2.600e-04 2.400e-04 2.200e-04 2.100e-04 1.900e-04
## Cumulative Proportion  9.967e-01 9.969e-01 9.972e-01 9.974e-01 9.976e-01
##                             PC47      PC48      PC49      PC50      PC51
## Standard deviation     7.946e-05 7.675e-05 7.607e-05 7.534e-05 7.418e-05
## Proportion of Variance 1.800e-04 1.700e-04 1.700e-04 1.700e-04 1.600e-04
## Cumulative Proportion  9.978e-01 9.979e-01 9.981e-01 9.983e-01 9.984e-01
##                             PC52      PC53      PC54      PC55      PC56
## Standard deviation     6.831e-05 0.0000662 6.387e-05 6.198e-05 5.748e-05
## Proportion of Variance 1.400e-04 0.0001300 1.200e-04 1.100e-04 1.000e-04
## Cumulative Proportion  9.986e-01 0.9986800 9.988e-01 9.989e-01 9.990e-01
##                             PC57      PC58      PC59      PC60      PC61
## Standard deviation     5.676e-05 5.547e-05 5.454e-05 5.088e-05 4.965e-05
## Proportion of Variance 9.000e-05 9.000e-05 9.000e-05 8.000e-05 7.000e-05
## Cumulative Proportion  9.991e-01 9.992e-01 9.993e-01 9.994e-01 9.994e-01
##                             PC62      PC63      PC64      PC65      PC66
## Standard deviation     4.769e-05 4.712e-05 4.393e-05 4.341e-05 4.172e-05
## Proportion of Variance 7.000e-05 6.000e-05 6.000e-05 5.000e-05 5.000e-05
## Cumulative Proportion  9.995e-01 9.996e-01 9.996e-01 9.997e-01 9.997e-01
##                             PC67      PC68      PC69      PC70      PC71
## Standard deviation     0.0000383 3.321e-05 0.0000328 3.133e-05 2.967e-05
## Proportion of Variance 0.0000400 3.000e-05 0.0000300 3.000e-05 3.000e-05
## Cumulative Proportion  0.9997700 9.998e-01 0.9998300 9.999e-01 9.999e-01
##                             PC72      PC73      PC74      PC75      PC76
## Standard deviation     2.761e-05 2.621e-05 2.522e-05 2.469e-05 1.898e-05
```

```
## Proportion of Variance 2.000e-05 2.000e-05 2.000e-05 2.000e-05 1.000e-05
## Cumulative Proportion  9.999e-01 9.999e-01 9.999e-01 1.000e+00 1.000e+00
##                             PC77      PC78      PC79      PC80      PC81
## Standard deviation     1.724e-05 1.681e-05 1.436e-05 1.146e-05 3.254e-06
## Proportion of Variance 1.000e-05 1.000e-05 1.000e-05 0.000e+00 0.000e+00
## Cumulative Proportion  1.000e+00 1.000e+00 1.000e+00 1.000e+00 1.000e+00
##                            PC82
## Standard deviation     1.04e-06
## Proportion of Variance 0.00e+00
## Cumulative Proportion  1.00e+00
```

AS can be seen in the skree plot above, the elbow of the plot is at 2. This means that there are 2 componants in the high volatility periods that can explain the sources of returns.