FA542 - Homework #1

I pledge my honor that I have abided by the Stevens Honor System.

Sid Bhatia

2023-09-16

Problem #1

Import Libraries

```
# Load libraries for skewness, kurtosis, and plotting.
library(moments)
library(fBasics)
```

```
##
## Attaching package: 'fBasics'

## The following objects are masked from 'package:moments':
##
##     kurtosis, skewness
```

Data Retrieval

```
# Establish the directory for data.
data_directory <- "C:/Users/sbhatia2/My Drive/University/Academics/Semester V/FA542 - Time Series with 

# Load in each dataset.
data_problem_1 <- read.table(paste(data_directory, 'HW1_1.txt', sep=""), header = T)
data_problem_2 <- read.table(paste(data_directory, 'HW1_2.txt', sep=""), header = T)
data_problem_3 <- read.table(paste(data_directory, 'HW1_3.txt', sep=""), header = T)
```

```

df_1 <- as.data.frame(data_problem_1)

# Create function that calculates the sample mean, standard deviation, skewness, kurtosis, minimum, and
compute_statistics <- function(returns)
{
  mean_val <- mean(returns)
  sd_val <- sd(returns)
  skewness_val <- skewness(returns)
  kurtosis_val <- kurtosis(returns)
  min_val <- min(returns)
  max_val <- max(returns)

  # Create a list to hold each sample statistic.
  result <- list(
    mean = mean_val,
```

```r
    standard_deviation = sd_val,
    skewness = skewness_val,
    kurtosis = kurtosis_val,
    minimum = min_val,
    maximum = max_val
  )

  return(result)
}

# Compute statistics for each simple return series.
CAT_statistics <- compute_statistics(df_1$RET)
VW_statistics <- compute_statistics(df_1$vwretd)
EW_statistics <- compute_statistics(df_1$ewretd)
SP_statistics <- compute_statistics(df_1$sprtrn)

CAT_statistics
```

1a)

```
## $mean
## [1] 0.0004945544
##
## $standard_deviation
## [1] 0.02092881
##
## $skewness
## [1] 0.2304287
## attr(,"method")
## [1] "moment"
##
## $kurtosis
## [1] 5.029022
## attr(,"method")
## [1] "excess"
##
## $minimum
## [1] -0.11434
##
## $maximum
## [1] 0.147229
```

```r
VW_statistics
```

```
## $mean
## [1] 0.0003372415
##
## $standard_deviation
## [1] 0.01320677
##
## $skewness
## [1] -0.1879131
## attr(,"method")
## [1] "moment"
##
```

```
## $kurtosis
## [1] 9.14096
## attr(,"method")
## [1] "excess"
##
## $minimum
## [1] -0.089771
##
## $maximum
## [1] 0.114887
```

```
EW_statistics
```

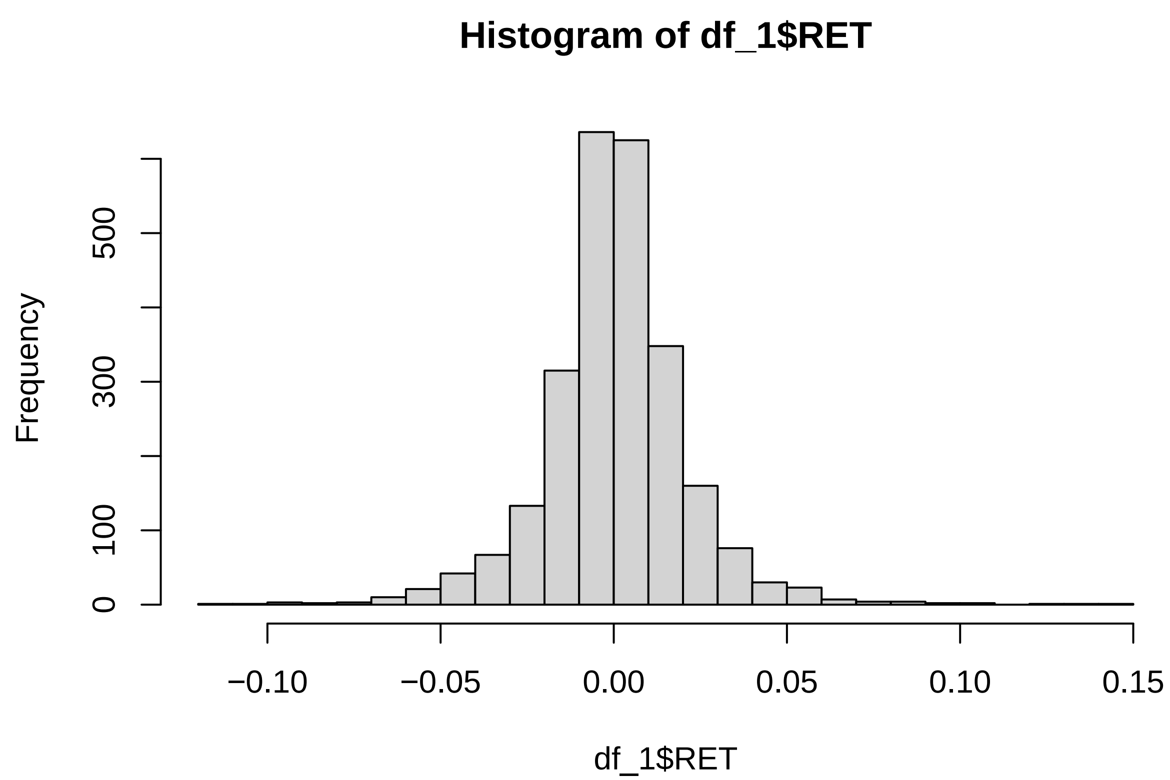
```
## $mean
## [1] 0.0004484198
##
## $standard_deviation
## [1] 0.01220235
##
## $skewness
## [1] -0.2246611
## attr(,"method")
## [1] "moment"
##
## $kurtosis
## [1] 7.847616
## attr(,"method")
## [1] "excess"
##
## $minimum
## [1] -0.07824
##
## $maximum
## [1] 0.107422
```

```
SP_statistics
```

```
## $mean
## [1] 0.0002682593
##
## $standard_deviation
## [1] 0.01317902
##
## $skewness
## [1] -0.08946497
## attr(,"method")
## [1] "moment"
##
## $kurtosis
## [1] 10.14406
## attr(,"method")
## [1] "excess"
##
## $minimum
## [1] -0.09035
##
```

```
## $maximum
## [1] 0.1158
```

```
# Display the empirical density distribution of simple returns for CAT.
hist(df_1$RET, nclass=30)
```



1b)

```
CAT_density_estimate <- density(df_1$RET)

# Obtain density estimate of simple returns for CAT.
CAT_density_estimate
```

```
##
## Call:
##  density.default(x = df_1$RET)
##
## Data: df_1$RET (2518 obs.);  Bandwidth 'bw' = 0.002757
##
##        x                  y
##  Min.   :-0.12261   Min.   : 0.00068
##  1st Qu.:-0.05308   1st Qu.: 0.06048
##  Median : 0.01644   Median : 0.22439
##  Mean   : 0.01644   Mean   : 3.59220
##  3rd Qu.: 0.08597   3rd Qu.: 2.89997
##  Max.   : 0.15550   Max.   :27.50193
```

```r
# Perform a Jarque-Bera test for normality of simple returns.
normalTest(df_1$RET, method = 'jb')
```

```
##
## Title:
##  Jarque - Bera Normalality Test
##
## Test Results:
##   STATISTIC:
##     X-squared: 2682.5018
##   P VALUE:
##     Asymptotic p Value: < 2.2e-16
```

At $\alpha = 0.05$ or at the 5% significance level, we reject the null hypothesis:

$$H_0 : \text{The simple returns of CAT are normally distributed}$$

since the p-value is less than 0.05.

1c) Log returns in relation to simple returns are defined as the following:

$$r_t = \ln(1 + R_t)$$

where r_t are log returns and R_t are simple returns.

```r
# Create function that transforms simple returns to log returns. log(...) is base e.
simple_to_log <- function(returns) {
  return(log(1 + returns))
}

# Convert simple returns to log using function.
CAT_log_returns <- simple_to_log(df_1$RET)
VW_log_returns <- simple_to_log(df_1$vwretd)
EW_log_returns <- simple_to_log(df_1$ewretd)
SP_log_returns <- simple_to_log(df_1$sprtrn)

# Use `compute_statistics()` function to calculate all relevant statistics.
CAT_log_statistics <- compute_statistics(CAT_log_returns)
VW_log_statistics <- compute_statistics(VW_log_returns)
EW_log_statistics <- compute_statistics(EW_log_returns)
SP_log_statistics <- compute_statistics(SP_log_returns)

CAT_log_statistics
```

```
## $mean
## [1] 0.0002760543
##
## $standard_deviation
## [1] 0.02089984
##
## $skewness
## [1] 0.01646851
## attr(,"method")
## [1] "moment"
##
```

```
## $kurtosis
## [1] 4.739097
## attr(,"method")
## [1] "excess"
##
## $minimum
## [1] -0.1214221
##
## $maximum
## [1] 0.1373495
```

```
VW_log_statistics
```

```
## $mean
## [1] 0.0002498325
##
## $standard_deviation
## [1] 0.01323116
##
## $skewness
## [1] -0.4052208
## attr(,"method")
## [1] "moment"
##
## $kurtosis
## [1] 9.027875
## attr(,"method")
## [1] "excess"
##
## $minimum
## [1] -0.09405906
##
## $maximum
## [1] 0.1087531
```

```
EW_log_statistics
```

```
## $mean
## [1] 0.0003737711
##
## $standard_deviation
## [1] 0.01222235
##
## $skewness
## [1] -0.4018266
## attr(,"method")
## [1] "moment"
##
## $kurtosis
## [1] 7.747635
## attr(,"method")
## [1] "excess"
##
## $minimum
## [1] -0.08147039
##
```

```
## $maximum
## [1] 0.1020348
```

```r
SP_log_statistics
```

```
## $mean
## [1] 0.0001812947
##
## $standard_deviation
## [1] 0.01319662
##
## $skewness
## [1] -0.3254499
## attr(,"method")
## [1] "moment"
##
## $kurtosis
## [1] 9.905808
## attr(,"method")
## [1] "excess"
##
## $minimum
## [1] -0.09469537
##
## $maximum
## [1] 0.1095716
```

1d) To test the following the null hypothesis that the mean log returns are zero, we need to conduct a t-test:

$$H_0 : \text{Mean log returns are zero.}$$

$$H_0 : \text{Mean log returns do not equal zero.}$$

```r
t.test(x = CAT_log_returns, alternative = c('two.sided'), mu = 0)
```

```
##
##  One Sample t-test
##
## data:  CAT_log_returns
## t = 0.6628, df = 2517, p-value = 0.5075
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.0005406637  0.0010927723
## sample estimates:
##    mean of x
## 0.0002760543
```

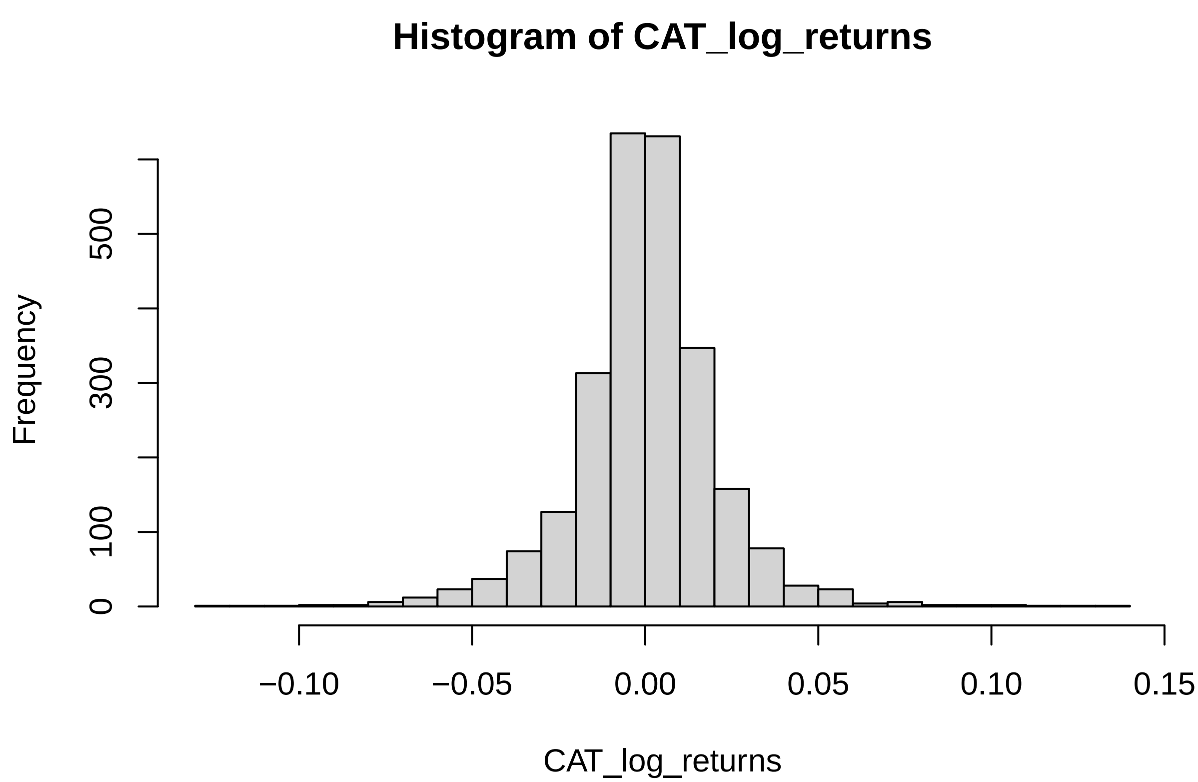
```r
t.test(x = SP_log_returns, alternative = c('two.sided'), mu = 0)
```

```
##
##  One Sample t-test
##
## data:  SP_log_returns
```

```
## t = 0.68937, df = 2517, p-value = 0.4907
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.0003343990  0.0006969884
## sample estimates:
##    mean of x
## 0.0001812947
```

As seen above, at the $\alpha = 0.05$ or the 5% significance level, we fail to reject the null hypothesis that the mean log returns for both CAT and S&P are zero since the p-values are well above 0.05.

```r
# Display the empirical density distribution of log returns for CAT.
hist(CAT_log_returns, nclass=30)
```



1e)

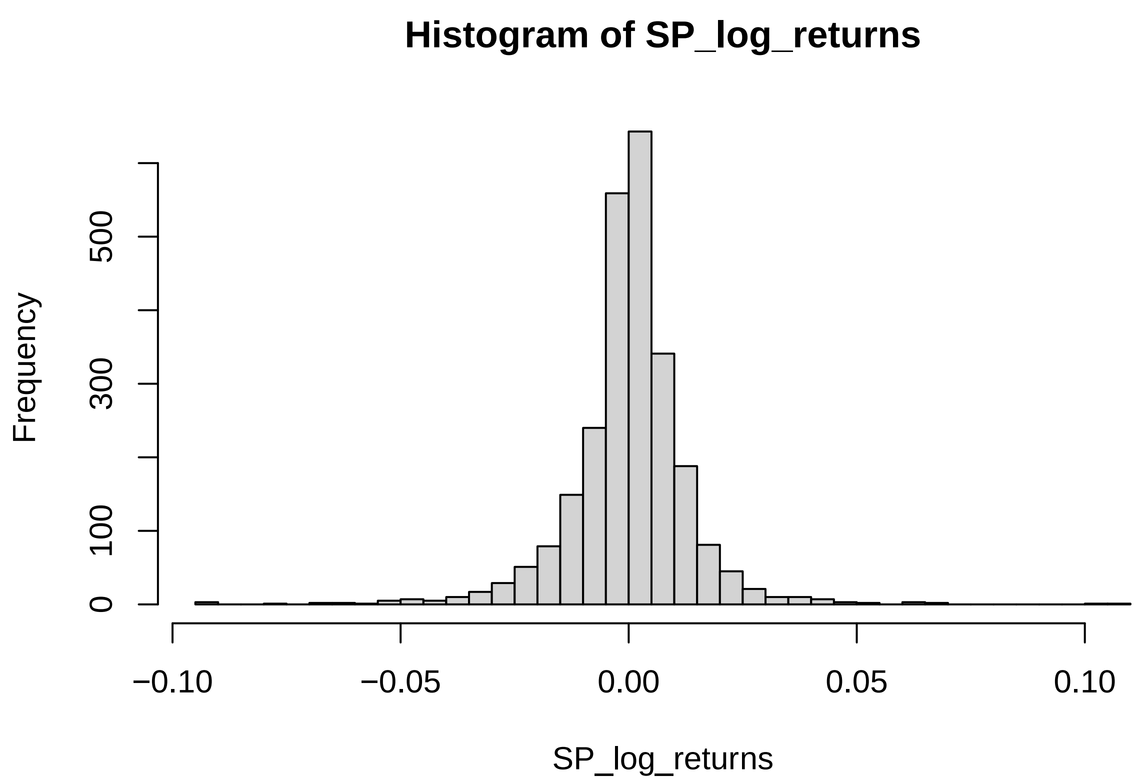
```r
CAT_log_density_estimate <- density(CAT_log_returns)

# Obtain density estimate of simple returns for CAT.
CAT_log_density_estimate
```

```
##
## Call:
##  density.default(x = CAT_log_returns)
##
## Data: CAT_log_returns (2518 obs.);   Bandwidth 'bw' = 0.002755
##
##        x                   y
```

```
##  Min.   :-0.129687   Min.   : 0.000663
##  1st Qu.:-0.060862   1st Qu.: 0.058853
##  Median : 0.007964   Median : 0.255782
##  Mean   : 0.007964   Mean   : 3.628847
##  3rd Qu.: 0.076789   3rd Qu.: 2.927840
##  Max.   : 0.145615   Max.   :27.510088
```

```
# Display the empirical density distribution of log returns for S&P
hist(SP_log_returns, nclass=30)
```




```
SP_log_density_estimate <- density(SP_log_returns)

# Obtain density estimate of simple returns for CAT.
SP_log_density_estimate
```

```
##
## Call:
##  density.default(x = SP_log_returns)
##
## Data: SP_log_returns (2518 obs.);    Bandwidth 'bw' = 0.001465
##
##        x                   y
##  Min.   :-0.099090   Min.   : 0.00000
##  1st Qu.:-0.045826   1st Qu.: 0.03419
##  Median : 0.007438   Median : 0.22369
##  Mean   : 0.007438   Mean   : 4.68935
##  3rd Qu.: 0.060702   3rd Qu.: 2.07732
```

```
##  Max.   : 0.113966   Max.   :55.73129
```

Problem #2

```r
df_2 <- as.data.frame(data_problem_2)

compute_statistics(df_2$RET)
```

1a)

```
## $mean
## [1] 0.01034235
##
## $standard_deviation
## [1] 0.05538326
##
## $skewness
## [1] -0.2969337
## attr(,"method")
## [1] "moment"
##
## $kurtosis
## [1] 3.220276
## attr(,"method")
## [1] "excess"
##
## $minimum
## [1] -0.357041
##
## $maximum
## [1] 0.250931
```

```r
compute_statistics(df_2$vwretd)
```

```
## $mean
## [1] 0.00887971
##
## $standard_deviation
## [1] 0.04403532
##
## $skewness
## [1] -0.522434
## attr(,"method")
## [1] "moment"
##
## $kurtosis
## [1] 1.984566
## attr(,"method")
## [1] "excess"
##
## $minimum
## [1] -0.225363
##
## $maximum
## [1] 0.165585
```

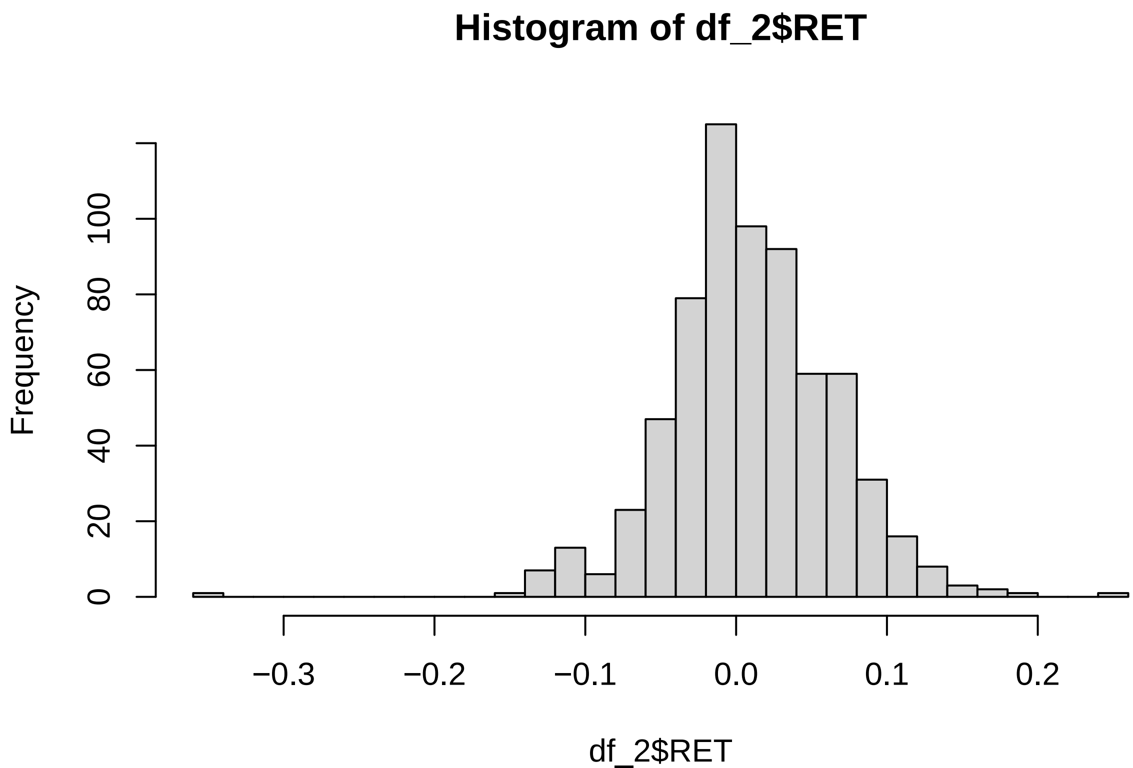
```r
compute_statistics(df_2$ewretd)
```

```
## $mean
## [1] 0.01136855
##
## $standard_deviation
## [1] 0.05554414
##
## $skewness
## [1] -0.1806713
## attr(,"method")
## [1] "moment"
##
## $kurtosis
## [1] 2.919352
## attr(,"method")
## [1] "excess"
##
## $minimum
## [1] -0.272248
##
## $maximum
## [1] 0.29926
```

```r
compute_statistics(df_2$sprtrn)
```

```
## $mean
## [1] 0.006359545
##
## $standard_deviation
## [1] 0.04251519
##
## $skewness
## [1] -0.4228902
## attr(,"method")
## [1] "moment"
##
## $kurtosis
## [1] 1.813676
## attr(,"method")
## [1] "excess"
##
## $minimum
## [1] -0.21763
##
## $maximum
## [1] 0.163047
```

```r
# Display the empirical density distribution of simple returns for PG.
hist(df_2$RET, nclass=30)
```



2b)

```
PG_density_estimate <- density(df_2$RET)

# Obtain density estimate of simple returns for CAT.
PG_density_estimate
```

```
##
## Call:
##  density.default(x = df_2$RET)
##
## Data: df_2$RET (672 obs.);   Bandwidth 'bw' = 0.01197
##
##        x                  y
##  Min.   :-0.39296   Min.   :0.000000
##  1st Qu.:-0.22301   1st Qu.:0.004268
##  Median :-0.05305   Median :0.079307
##  Mean   :-0.05305   Mean   :1.469559
##  3rd Qu.: 0.11690   3rd Qu.:1.647373
##  Max.   : 0.28685   Max.   :8.486108
```

```
# Perform a Jarque-Bera test for normality of simple returns.
normalTest(df_2$RET, method = 'jb')
```

```
##
## Title:
##  Jarque - Bera Normalality Test
##
## Test Results:
```

```
##   STATISTIC:
##     X-squared: 303.6398
##   P VALUE:
##     Asymptotic p Value: < 2.2e-16
```

At $\alpha = 0.05$ or at the 5% significance level, we reject the null hypothesis:

$$H_0 : \text{The simple returns of PG are normally distributed}$$

since the p-value is less than 0.05.

```
compute_statistics(simple_to_log(df_2$RET))
```

2c)

```
## $mean
## [1] 0.008756106
##
## $standard_deviation
## [1] 0.05580647
##
## $skewness
## [1] -0.8330476
## attr(,"method")
## [1] "moment"
##
## $kurtosis
## [1] 6.351822
## attr(,"method")
## [1] "excess"
##
## $minimum
## [1] -0.4416743
##
## $maximum
## [1] 0.2238881
```

```
compute_statistics(simple_to_log(df_2$vwretd))
```

```
## $mean
## [1] 0.007869959
##
## $standard_deviation
## [1] 0.04434013
##
## $skewness
## [1] -0.7880992
## attr(,"method")
## [1] "moment"
##
## $kurtosis
## [1] 2.859168
## attr(,"method")
## [1] "excess"
##
## $minimum
```

```
## [1] -0.2553607
##
## $maximum
## [1] 0.1532231
```

```
compute_statistics(simple_to_log(df_2$ewretd))
```

```
## $mean
## [1] 0.009774514
##
## $standard_deviation
## [1] 0.05564156
##
## $skewness
## [1] -0.5980258
## attr(,"method")
## [1] "moment"
##
## $kurtosis
## [1] 3.53031
## attr(,"method")
## [1] "excess"
##
## $minimum
## [1] -0.3177949
##
## $maximum
## [1] 0.2617949
```

```
compute_statistics(simple_to_log(df_2$sprtrn))
```

```
## $mean
## [1] 0.005433599
##
## $standard_deviation
## [1] 0.0427925
##
## $skewness
## [1] -0.6717913
## attr(,"method")
## [1] "moment"
##
## $kurtosis
## [1] 2.538153
## attr(,"method")
## [1] "excess"
##
## $minimum
## [1] -0.2454275
##
## $maximum
## [1] 0.1510433
```

2d) To test the following the null hypothesis that the mean log returns are zero, we need to conduct a t-test:

$$H_0 : \text{Mean log returns are zero.}$$

$$H_0 : \text{Mean log returns do not equal zero.}$$

```
t.test(x = simple_to_log(df_2$RET), alternative = c('two.sided'), mu = 0)
```

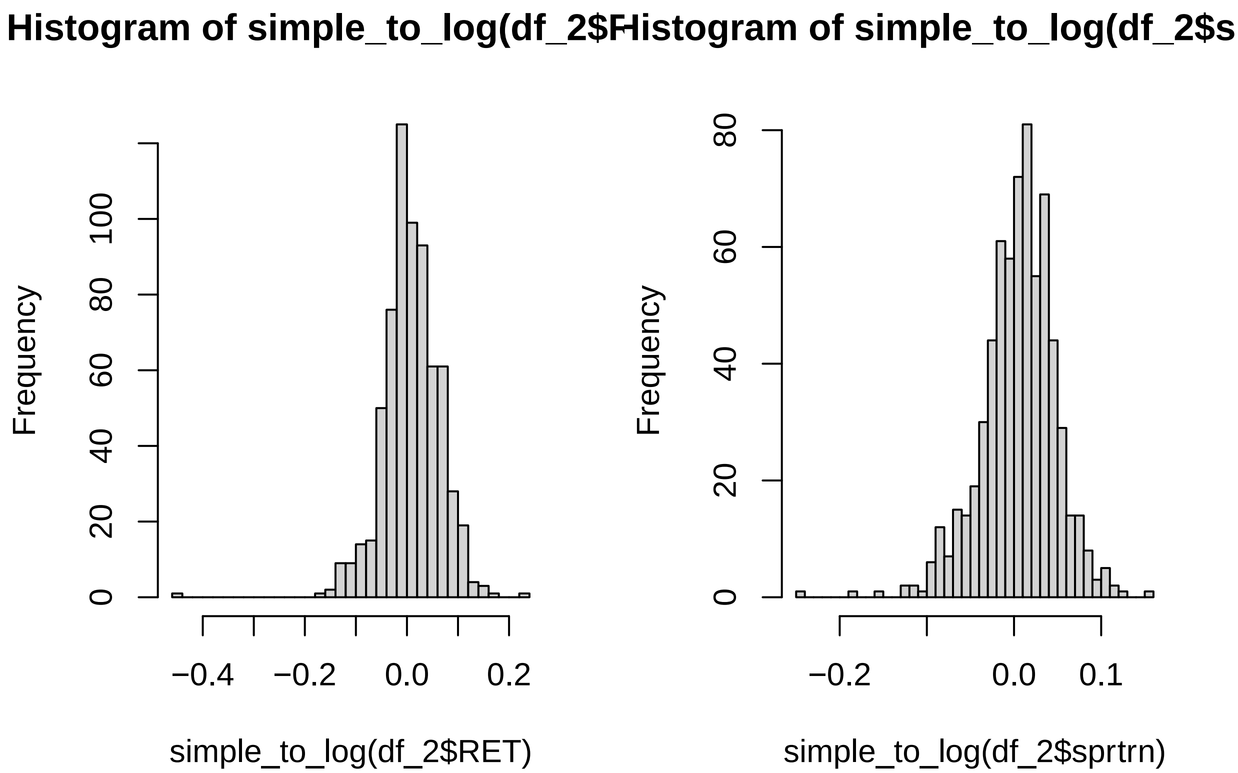
```
##
##  One Sample t-test
##
## data:  simple_to_log(df_2$RET)
## t = 4.0673, df = 671, p-value = 5.32e-05
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.004529107 0.012983105
## sample estimates:
##   mean of x
## 0.008756106
```

```
t.test(x = simple_to_log(df_2$sprtrn), alternative = c('two.sided'), mu = 0)
```

```
##
##  One Sample t-test
##
## data:  simple_to_log(df_2$sprtrn)
## t = 3.2916, df = 671, p-value = 0.001048
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.002192330 0.008674869
## sample estimates:
##   mean of x
## 0.005433599
```

As seen above, at the $\alpha = 0.05$ or the 5% significance level, we reject the null hypothesis that the mean log returns for both CAT and S&P are zero since the p-values below 0.05.

```
par(mfrow = c(1, 2))
hist(simple_to_log(df_2$RET), nclass=30)
hist(simple_to_log(df_2$sprtrn), nclass=30)
```



2e)

```
density(simple_to_log(df_2$RET))
```

```
##
## Call:
##  density.default(x = simple_to_log(df_2$RET))
##
## Data: simple_to_log(df_2$RET) (672 obs.);    Bandwidth 'bw' = 0.01184
##
##        x                  y
##  Min.   :-0.47720   Min.   :0.000000
##  1st Qu.:-0.29305   1st Qu.:0.000035
##  Median :-0.10889   Median :0.049018
##  Mean   :-0.10889   Mean   :1.356242
##  3rd Qu.: 0.07526   3rd Qu.:1.360786
##  Max.   : 0.25941   Max.   :8.473655
```

```
density(simple_to_log(df_2$sprtrn))
```

```
##
## Call:
##  density.default(x = simple_to_log(df_2$sprtrn))
##
## Data: simple_to_log(df_2$sprtrn) (672 obs.); Bandwidth 'bw' = 0.009471
##
##        x                  y
##  Min.   :-0.27384   Min.   : 0.000711
##  1st Qu.:-0.16052   1st Qu.: 0.050301
```

```
## Median :-0.04719   Median : 0.352586
## Mean   :-0.04719   Mean   : 2.203906
## 3rd Qu.: 0.06613   3rd Qu.: 2.879510
## Max.   : 0.17946   Max.   :10.731430
```

Problem 3