Live Case: S&P500 (2)

Aug 7, 2023

```
# Load the required libraries, suppressing annoying startup messages
library(dplyr, quietly = TRUE, warn.conflicts = FALSE)
library(tibble, quietly = TRUE, warn.conflicts = FALSE)
library(ggplot2, quietly = TRUE, warn.conflicts = FALSE) # For data visualization
library(ggpubr, quietly = TRUE, warn.conflicts = FALSE) # For data visualization

library(gsheet, quietly = TRUE, warn.conflicts = FALSE)
library(rmarkdown, quietly = TRUE, warn.conflicts = FALSE)
library(knitr, quietly = TRUE, warn.conflicts = FALSE)
library(kableExtra, quietly = TRUE, warn.conflicts = FALSE)
```

ISSUE: Understanding the S&P500 as a whole *Aug 06, 2023 -=- This chapter is being heavily edited*

S&P 500 Data - PRELIMINARY SETUP

1. We will continue our analysis of the S&P 500. Load the data, as described in the chapter Live Case: S&P500 (1 of 3)

```
# Read S&P500 stock data present in a Google Sheet.
library(gsheet)
prefix <- "https://docs.google.com/spreadsheets/d/"
sheetID <- "11ahk9uWxBkDqrhNm7qYmiTwrlSC53N1zvXYfv7ttOCM"
url500 <- paste(prefix,sheetID) # Form the URL to connect to
sp500 <- gsheet2tbl(url500) # Read it into a tibble called sp500
```

2. Rename columns, as described in the chapter Live Case: S&P500 (1 of 3).

```
suppressPackageStartupMessages(library(dplyr))
```

```r
# Define a mapping of new column names
new_names <- c(
  "Date", "Stock", "StockName", "Sector", "Industry",
  "MarketCap", "Price", "Low52Wk", "High52Wk",
  "ROE", "ROA", "ROIC", "GrossMargin",
  "OperatingMargin", "NetMargin", "PE",
  "PB", "EVEBITDA", "EBITDA", "EPS",
  "EBITDA_YOY", "EBITDA_QYOY", "EPS_YOY",
  "EPS_QYOY", "PFCF", "FCF",
  "FCF_QYOY", "DebtToEquity", "CurrentRatio",
  "QuickRatio", "DividendYield",
  "DividendsPerShare_YOY", "PS",
  "Revenue_YOY", "Revenue_QYOY", "Rating"
)
# Rename the columns using the new_names vector
sp500 <- sp500 %>%
  rename_with(~ new_names, everything())
```

3. Remove Rows containing no data or Null values, as described in the chapter Live Case: S&P500 (1 of 3).

```r
# Check for blank or null values in the "Stock" column
hasNull <- any(sp500$Stock == "" | is.null(sp500$Stock))
if (hasNull) {
    # Remove rows with null or blank values from the dataframe tibble
    sp500 <- sp500[!(is.null(sp500$Stock) | sp500$Stock == ""), ]
}
```

4. The S&P500 shares are divided into multiple Sectors. Thus, model Sector as a factor() variable, as described in the chapter Live Case: S&P500 (1 of 3).

```r
sp500$Sector <- as.factor(sp500$Sector)
```

5. Stock Ratings: The S&P500 shares have Technical Ratings such as {Buy, Sell, ..}. Model the data column Rating as a factor() variable, as described in the chapter Live Case: S&P500 (1 of 3).

```r
sp500$Rating <- as.factor(sp500$Rating)
```

6. Low52WkPerc: Create a new column to track Share Prices relative to their 52 Week Low, as described in the chapter Live Case: S&P500 (1 of 3).

```
sp500 <- sp500 %>% mutate(Low52WkPerc = round((Price - Low52Wk)*100 / Low52Wk,2))
colnames(sp500)
```

```
 [1] "Date"                "Stock"               "StockName"
 [4] "Sector"              "Industry"            "MarketCap"
 [7] "Price"               "Low52Wk"             "High52Wk"
[10] "ROE"                 "ROA"                 "ROIC"
[13] "GrossMargin"         "OperatingMargin"     "NetMargin"
[16] "PE"                  "PB"                  "EVEBITDA"
[19] "EBITDA"              "EPS"                 "EBITDA_YOY"
[22] "EBITDA_QYOY"         "EPS_YOY"             "EPS_QYOY"
[25] "PFCF"                "FCF"                 "FCF_QYOY"
[28] "DebtToEquity"        "CurrentRatio"        "QuickRatio"
[31] "DividendYield"       "DividendsPerShare_YOY" "PS"
[34] "Revenue_YOY"         "Revenue_QYOY"        "Rating"
[37] "Low52WkPerc"
```

Well done! Our data is now ready for analysis!!

7. Low52WkPerc: Create a new column MarketCapBillions = MarketCap/1000,000,000, as described in the chapter Live Case: S&P500 (1 of 3).

```
sp500 <- sp500 %>% mutate(MarketCapBillions = round(MarketCap/1000000000))
colnames(sp500)
```

```
 [1] "Date"                "Stock"               "StockName"
 [4] "Sector"              "Industry"            "MarketCap"
 [7] "Price"               "Low52Wk"             "High52Wk"
[10] "ROE"                 "ROA"                 "ROIC"
[13] "GrossMargin"         "OperatingMargin"     "NetMargin"
[16] "PE"                  "PB"                  "EVEBITDA"
[19] "EBITDA"              "EPS"                 "EBITDA_YOY"
[22] "EBITDA_QYOY"         "EPS_YOY"             "EPS_QYOY"
[25] "PFCF"                "FCF"                 "FCF_QYOY"
[28] "DebtToEquity"        "CurrentRatio"        "QuickRatio"
[31] "DividendYield"       "DividendsPerShare_YOY" "PS"
[34] "Revenue_YOY"         "Revenue_QYOY"        "Rating"
[37] "Low52WkPerc"         "MarketCapBillions"
```

ANALYSIS OF S&P500 SECTORS

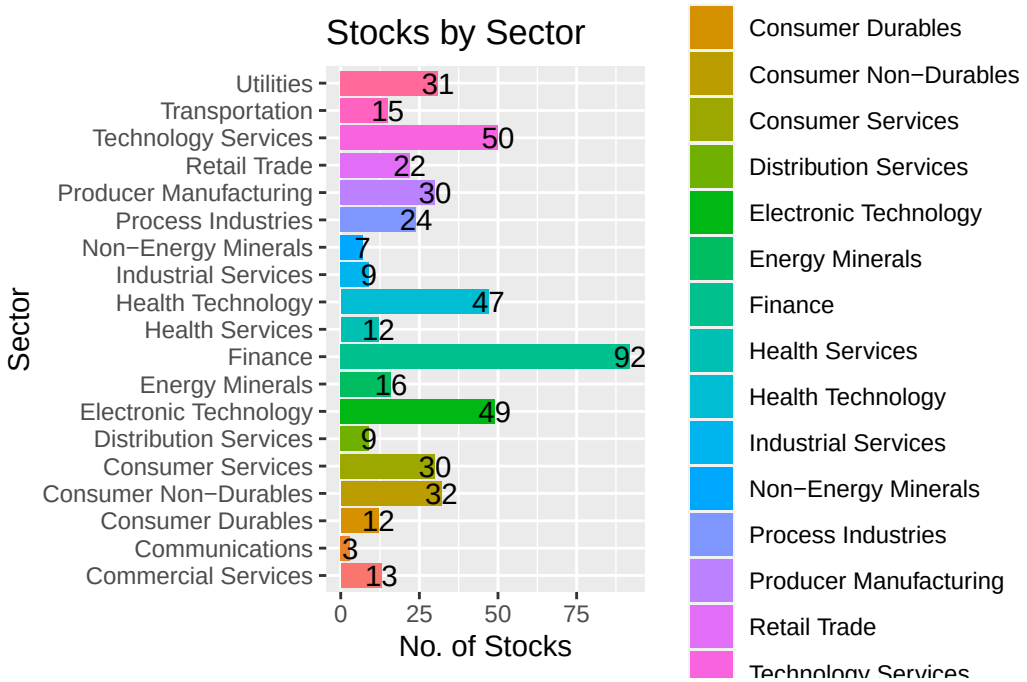
- The table() function allows us to count how many stocks are part of each sector.

```
tab<- addmargins(table(sp500$Sector))
kable(tab)
```

Var1	Freq
Commercial Services	13
Communications	3
Consumer Durables	12
Consumer Non-Durables	32
Consumer Services	30
Distribution Services	9
Electronic Technology	49
Energy Minerals	16
Finance	92
Health Services	12
Health Technology	47
Industrial Services	9
Non-Energy Minerals	7
Process Industries	24
Producer Manufacturing	30
Retail Trade	22
Technology Services	50
Transportation	15
Utilities	31
Sum	503

- The S&P500 consists of 503 stocks, divided across 19 sectors.

```
ggplot(data = sp500,
       aes(y = Sector)) +
  geom_bar(aes(fill = Sector)) +
  geom_text(stat='count',
            aes(label=after_stat(count))) +
  labs(title = "Stocks by Sector",
       x = "No. of Stocks",
       y = "Sector")
```



- Numbers of shares by Rating

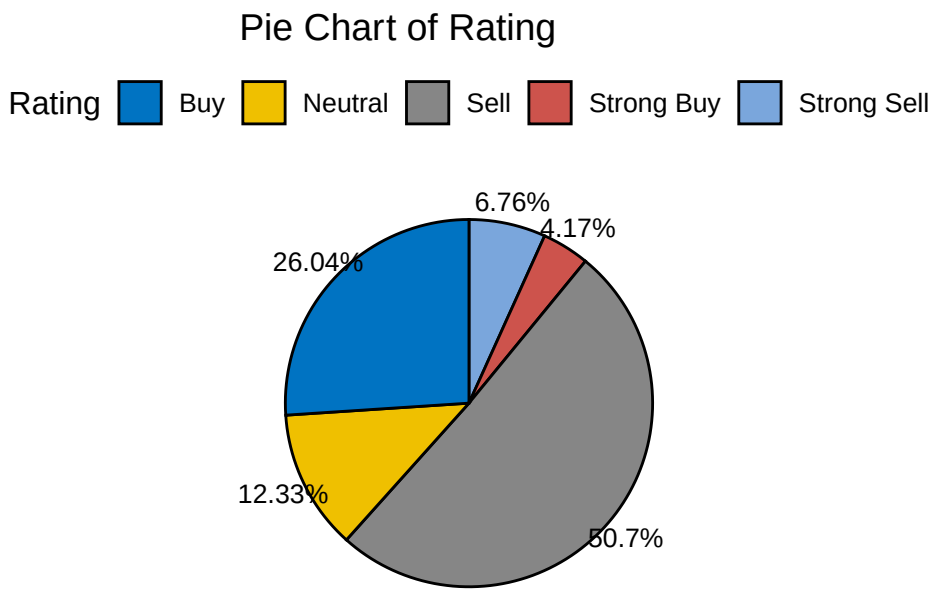
```
tab<- addmargins(table(sp500$Rating))
kable(tab)
```

Var1	Freq
Buy	131
Neutral	62
Sell	255
Strong Buy	21
Strong Sell	34
Sum	503

- Pie Chart Showing Proportion of shares by Rating

```
library(ggpubr)

# Compute counts and proportions of each cylinder type
Rating_counts <- as.data.frame(table(sp500$Rating))
colnames(Rating_counts) <- c("Rating", "n")

# Calculate proportions
```

```r
Rating_counts$prop <- Rating_counts$n / sum(Rating_counts$n)

# Create labels that display proportions as percentages
Rating_counts$labels <- paste0(round(Rating_counts$prop*100, 2), "%")

# Create the pie chart with proportions
ggpie(data = Rating_counts,
      x = "prop",
      fill = "Rating",
      label = "labels",
      palette = "jco",
      title = "Pie Chart of Rating")
```



- Count Shares by Sector*Rating

```r
tab<- addmargins(table(Sector = sp500$Sector, Rating = sp500$Rating))
kable(tab)
```

	Buy	Neutral	Sell	Strong Buy	Strong Sell	Sum
Commercial Services	2	1	8	1	1	13
Communications	0	1	1	0	1	3
Consumer Durables	1	3	7	1	0	12
Consumer Non-Durables	12	3	9	3	5	32
Consumer Services	6	2	17	3	2	30
Distribution Services	3	1	5	0	0	9
Electronic Technology	10	8	30	1	0	49
Energy Minerals	2	3	8	0	3	16
Finance	34	17	37	1	3	92
Health Services	2	0	6	0	4	12
Health Technology	11	5	27	1	3	47
Industrial Services	2	0	7	0	0	9
Non-Energy Minerals	2	1	4	0	0	7
Process Industries	5	5	11	1	2	24
Producer Manufacturing	5	2	18	2	3	30
Retail Trade	8	1	10	1	2	22
Technology Services	13	3	28	2	4	50
Transportation	4	2	7	1	1	15
Utilities	9	4	15	3	0	31
Sum	131	62	255	21	34	503

MARKET CAP

TODO: Work in Billions; 1. Market Cap of all companies by Sector

```
library(janitor)  # This package helps us auto generate the total at the bottom of a table
library(kableExtra)

# Market Cap by Sector
MCap <- sp500 %>%
  group_by(Sector) %>%
  summarise(
    MarketCapCr = sum(na.omit(MarketCap)/10000000))

# Total Market Cap of the entire S&P 500 (in Millions)
SP500MarketCap <- sum(sp500$MarketCap/10000000)

# calculating % market cap
PercentMarketCap <- round(MCap$MarketCapCr*100/SP500MarketCap,2)
MCapTab <- cbind(MCap,PercentMarketCap)
```

```r
# sorting by PercentMarketCap
MCapTab <- MCapTab %>% arrange(desc(PercentMarketCap))

# Use package janitor to add sums at the bottom of the table
MCapTab <- MCapTab %>%
  adorn_totals("row")

# Use package knittr to format the appearance of the table
MCapTab <- knitr::kable(MCapTab, "html") %>% kable_styling()
MCapTab
```

Sector	MarketCapCr	PercentMarketCap
Technology Services	925340.66	23.13
Electronic Technology	635429.49	15.88
Finance	474633.96	11.86
Health Technology	375238.19	9.38
Retail Trade	298136.12	7.45
Consumer Non-Durables	197559.68	4.94
Energy Minerals	145577.31	3.64
Consumer Services	141199.13	3.53
Producer Manufacturing	130414.37	3.26
Commercial Services	123592.53	3.09
Consumer Durables	104930.56	2.62
Health Services	92187.36	2.30
Utilities	89044.61	2.23
Process Industries	74642.37	1.87
Transportation	62170.55	1.55
Communications	40839.40	1.02
Industrial Services	40376.71	1.01
Distribution Services	30319.84	0.76
Non-Energy Minerals	19715.17	0.49
Total	4001348.01	100.01

TODO: Work in Billions; Show “Sum”, “Median”, “Mean” Delete Q1, Q3 all others 2. Summary Statistics of Market Cap (in Cr of USD) by each Sector of S&P500

```r
SectorMC <- sp500 %>%
  group_by(Sector) %>%
  summarise(
```

```
    Mean = mean(na.omit(MarketCap/10000000)),
    Median= sd(na.omit(MarketCap/10000000)),
    Median= median(na.omit(MarketCap/10000000)),
    Q1 = quantile(na.omit(MarketCap/10000000), probs = 0.25, na.rm = TRUE),
    Q3 = quantile(na.omit(MarketCap/10000000), probs = 0.75, na.rm = TRUE),
    Min = min(na.omit(MarketCap/10000000)),
    max = max(na.omit(MarketCap/10000000))
  )

tab <- cbind(Sector = SectorMC$Sector, round(SectorMC[,2:7],2))

SMcap <- knitr::kable(tab, "html") %>% kable_styling()
SMcap
```

Sector	Mean	Median	Q1	Q3	Min	max
Commercial Services	9507.12	3000.07	1471.51	6419.13	784.70	46815.30
Communications	13613.13	13625.30	12181.55	15050.80	10737.80	16476.30
Consumer Durables	8744.21	1785.81	1313.68	3860.84	732.91	79419.70
Consumer Non-Durables	6173.74	3782.36	1710.76	5629.61	687.13	34377.70
Consumer Services	4706.64	2135.78	1362.34	5660.04	701.10	19198.50
Distribution Services	3368.87	3121.76	2027.64	3614.77	969.59	7731.10
Electronic Technology	12967.95	3878.08	1726.29	7451.34	815.61	267674.00
Energy Minerals	9098.58	5177.19	2587.75	6383.61	1262.86	47069.50
Finance	5159.06	2413.20	1579.81	4951.02	516.88	76398.60
Health Services	7682.28	3666.32	1677.90	7134.55	781.29	46703.40
Health Technology	7983.79	3271.71	1660.28	10323.95	443.67	50989.50
Industrial Services	4486.30	3694.64	3639.11	4507.96	2716.24	8285.51
Non-Energy Minerals	2816.45	2684.16	2156.49	3412.77	546.46	5346.03
Process Industries	3110.10	1642.52	1300.52	3626.58	474.53	18168.70
Producer Manufacturing	4347.15	3271.40	1347.46	5126.48	678.20	13926.90
Retail Trade	13551.64	3521.89	2027.16	9868.37	658.75	130430.00
Technology Services	18506.81	3347.85	1671.89	11042.51	427.38	234595.00
Transportation	4144.70	2380.65	1499.87	5320.04	471.75	13318.30
Utilities	2872.41	2046.91	1633.70	3440.55	834.88	11750.50

3. Top 10 companies having highest Market Cap

```
Top10 <- sp500 %>% arrange(desc(MarketCap)) %>% head(10)
Top10 <- Top10[,c(1:4, 6,10:13)]
```

```
Top10 <- knitr::kable(Top10, "html") %>% kable_styling()
Top10
```

Date	Stock	StockName	Sector	MarketCap	ROE	ROA
9/30/2023	AAPL	Apple Inc.	Electronic Technology	2.67674e+12	160.1	28.2
9/30/2023	MSFT	Microsoft Corporation	Technology Services	2.34595e+12	38.8	18.6
9/30/2023	GOOG	Alphabet Inc.	Technology Services	1.65589e+12	23.3	16.5
9/30/2023	GOOGL	Alphabet Inc.	Technology Services	1.65589e+12	23.3	16.5
9/30/2023	AMZN	Amazon.com, Inc.	Retail Trade	1.30430e+12	8.7	2.9
9/30/2023	NVDA	NVIDIA Corporation	Electronic Technology	1.07443e+12	40.2	22.2
9/30/2023	TSLA	Tesla, Inc.	Consumer Durables	7.94197e+11	28.0	15.4
9/30/2023	META	Meta Platforms, Inc.	Technology Services	7.72489e+11	17.4	12.0
9/30/2023	BRK.B	Berkshire Hathaway Inc. New	Finance	7.63986e+11	17.4	8.9
9/30/2023	LLY	Eli Lilly and Company	Health Technology	5.09895e+11	66.3	12.8

PRICE RELATIVE TO 52 WEEK LOW

1. Summary Statistics of Low52WkPerc by Sector

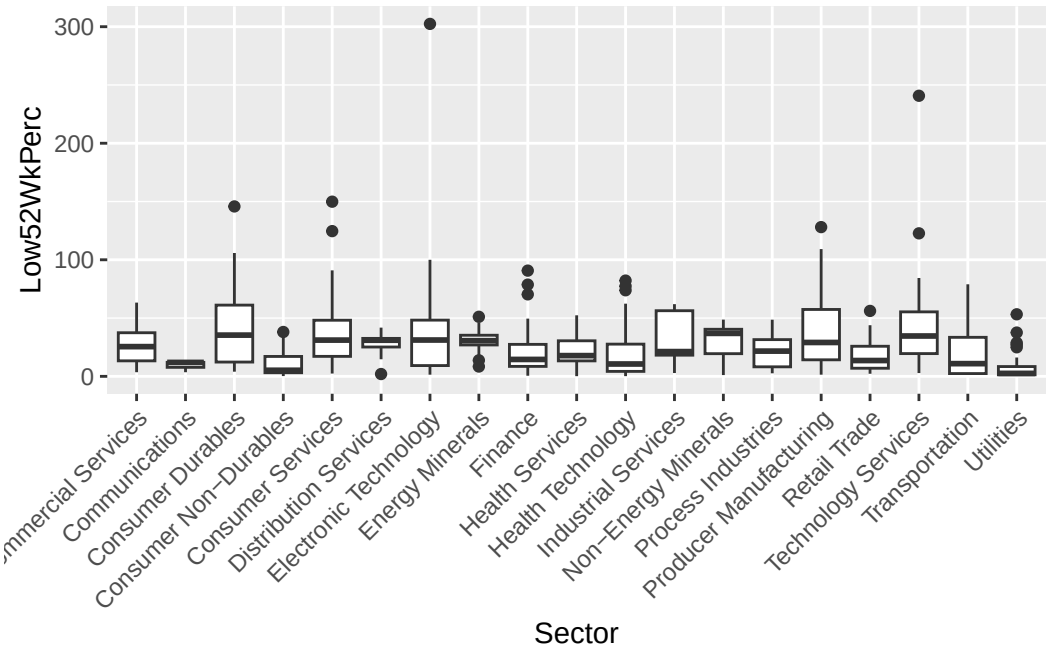
```
SM <- sp500 %>%
  group_by(Sector) %>%
  summarise(
    Mean = mean(na.omit(Low52WkPerc)),
    Median= sd(na.omit(Low52WkPerc)),
    Median= median(na.omit(Low52WkPerc)),
    Q1 = quantile(na.omit(Low52WkPerc), probs = 0.25, na.rm = TRUE),
    Q3 = quantile(na.omit(Low52WkPerc), probs = 0.75, na.rm = TRUE),
    Min = min(na.omit(Low52WkPerc)),
    Max = max(na.omit(Low52WkPerc))
  )

tab <- cbind(Sector = SM$Sector, round(SM[,2:7],2))

tab <- tab %>% arrange(Median)

SM <- knitr::kable(tab, "html") %>% kable_styling()
SM
```

Sector	Mean	Median	Q1	Q3	Min	Max
Utilities	8.72	2.54	1.33	8.39	0.53	53.23
Consumer Non-Durables	10.35	5.21	3.11	17.07	0.24	38.05
Health Technology	19.30	10.60	4.24	27.63	0.17	82.15
Transportation	21.23	10.89	2.38	33.48	1.49	78.99
Communications	9.18	11.94	7.72	12.02	3.51	12.09
Retail Trade	18.10	13.55	6.96	25.80	2.22	56.14
Finance	19.63	14.51	8.53	27.42	0.53	90.71
Health Services	22.08	17.77	13.27	30.50	0.16	52.36
Industrial Services	32.57	21.22	18.16	56.30	2.83	61.93
Process Industries	22.11	21.63	8.16	31.50	2.60	48.61
Commercial Services	28.58	25.48	13.29	37.40	3.54	63.26
Producer Manufacturing	39.46	29.06	14.18	57.38	1.41	128.04
Energy Minerals	30.49	30.60	26.93	35.26	8.41	51.10
Distribution Services	26.96	30.98	25.17	32.55	1.98	41.81
Consumer Services	39.73	31.10	17.14	48.15	2.40	149.86
Electronic Technology	40.13	31.16	9.24	48.23	1.41	302.41
Technology Services	42.46	34.62	19.51	55.33	2.82	240.75
Consumer Durables	46.04	35.37	12.26	61.11	3.99	145.78
Non-Energy Minerals	29.47	36.86	19.44	40.37	1.09	48.68

Sector Communications and Utilities are closest to its 52 week low.

2. Box Plot for Low52WkPerc by Sector TODO: Truncate at 100; Rotate by 90 degrees; Sort Sectors by Median(Low52WkPerc)

```
library(ggplot2)

ggplot(sp500, aes(Sector, Low52WkPerc)) + geom_boxplot() +
  theme(axis.text.x = element_text(angle = 45, vjust = 1, hjust = 1))
```



PROFITABILITY BY SECTOR

ROE

1. Summary Statistics of ROE by each Sector of S&P500

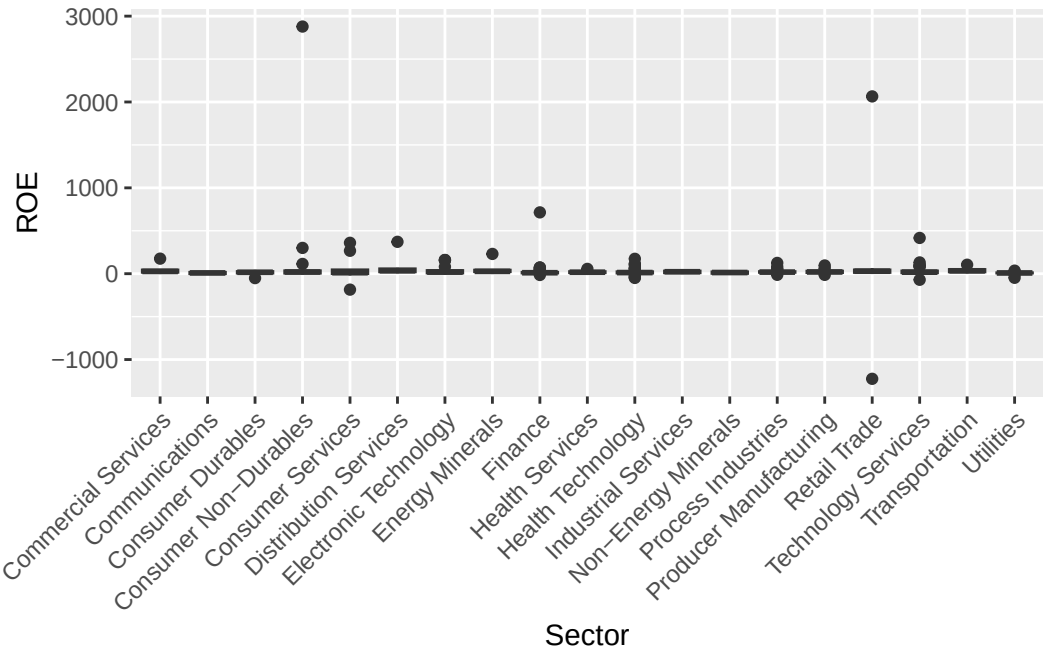
```
SectorROE <- sp500 %>%
  group_by(Sector) %>%
  summarise(
    Mean = mean(na.omit(ROE)),
    Median= sd(na.omit(ROE)),
    Median= median(na.omit(ROE)),
    Q1 = quantile(na.omit(ROE), probs = 0.25, na.rm = TRUE),
    Q3 = quantile(na.omit(ROE), probs = 0.75, na.rm = TRUE),
    Min = min(na.omit(ROE)),
    max = max(na.omit(ROE))
  )

cbind(Sector = SectorROE$Sector, round(SectorROE[,2:7],2))
```

```
                  Sector   Mean Median    Q1    Q3     Min    max
1     Commercial Services  37.98  26.40 16.40 43.60     3.5  175.2
```

```
2          Communications   8.10   9.10  0.55 16.15    -8.0   23.2
3       Consumer Durables  12.23  16.65  6.85 25.38   -51.4   45.2
4   Consumer Non-Durables 129.53  19.60  6.40 33.90   -11.5 2878.8
5       Consumer Services  33.02  11.40  1.55 44.95  -185.6  359.9
6   Distribution Services  81.10  34.20 22.15 56.45     5.1  371.2
7   Electronic Technology  31.51  18.75  8.10 36.80   -14.8  160.1
8         Energy Minerals  43.12  26.95 23.78 41.45    18.0  230.2
9                 Finance  22.13  11.00  7.82 16.67   -14.7  714.3
10        Health Services  20.63  17.30 12.05 24.05     8.3   56.0
11      Health Technology  19.87  13.10  6.80 22.73   -49.3  173.5
12    Industrial Services  21.04  22.60 10.70 31.10     7.7   36.5
13    Non-Energy Minerals  13.84  13.50  3.40 21.80    -3.8   36.8
14     Process Industries  25.72  18.60 15.35 24.62   -13.2  125.5
15 Producer Manufacturing  25.42  20.20 13.02 30.00   -13.6   95.9
16           Retail Trade  74.34  28.70 14.47 44.00 -1224.5 2065.3
17    Technology Services  33.17  18.00 10.70 31.82   -70.6  416.6
18         Transportation  36.39  33.50 20.85 49.08     4.1  104.4
19              Utilities   8.12   8.70  7.65 10.60   -47.6   35.5
```

2. Box Plot for ROE by Sector

```r
library(ggplot2)

ggplot(sp500, aes(Sector, ROE)) + geom_boxplot() +
  theme(axis.text.x = element_text(angle = 45, vjust = 1, hjust = 1))
```



ROA

1. Summary Statistics of ROA by each Sector of S&P500

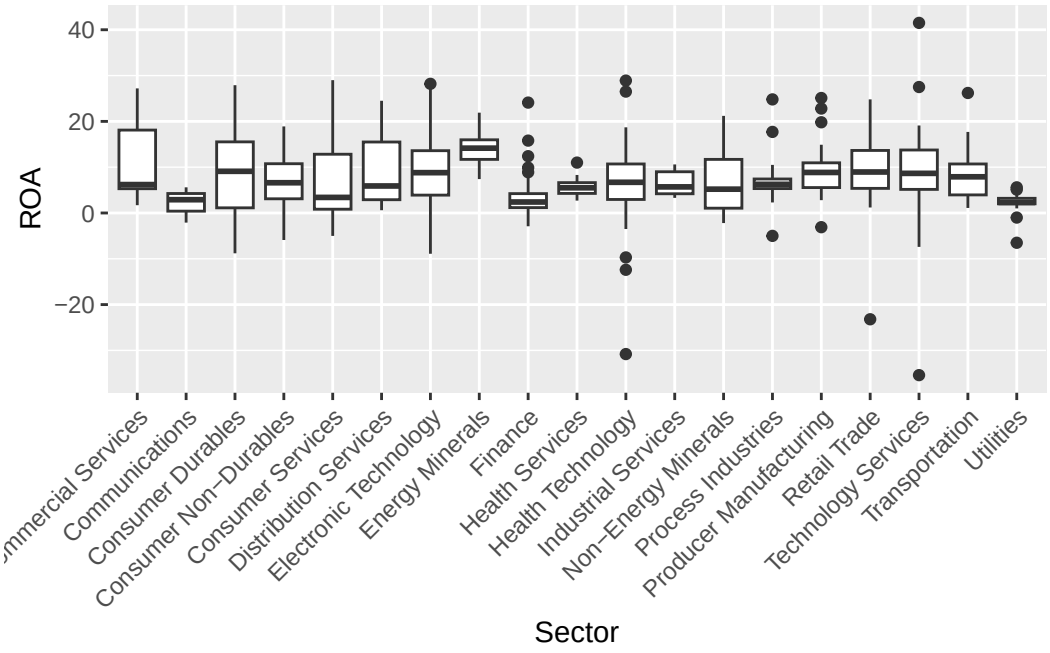
```r
SectorROA <- sp500 %>%
  group_by(Sector) %>%
  summarise(
    Mean = mean(na.omit(ROA)),
    Median= sd(na.omit(ROA)),
    Median= median(na.omit(ROA)),
    Q1 = quantile(na.omit(ROA), probs = 0.25, na.rm = TRUE),
    Q3 = quantile(na.omit(ROA), probs = 0.75, na.rm = TRUE),
    Min = min(na.omit(ROA)),
    max = max(na.omit(ROA))
  )

cbind(Sector = SectorROA$Sector, round(SectorROA[,2:7],2))
```

```
                  Sector  Mean Median    Q1    Q3   Min  max
1     Commercial Services 10.25   6.20  5.30 18.10   1.7 27.2
2          Communications  2.13   2.90  0.40  4.25  -2.1  5.6
3       Consumer Durables  8.65   9.10  1.13 15.53  -8.8 27.9
```

```
4   Consumer Non-Durables  7.14   6.60  3.10 10.75  -5.9 18.9
5       Consumer Services  6.71   3.40  0.83 12.83  -5.0 29.0
6   Distribution Services  9.79   5.90  2.90 15.50   0.6 24.5
7   Electronic Technology  9.37   8.80  3.90 13.60  -8.9 28.2
8         Energy Minerals 14.32  14.15 11.70 15.98   7.4 21.9
9                 Finance  3.32   2.40  1.17  4.23  -2.9 24.1
10        Health Services  5.74   5.50  4.27  6.62   2.7 11.0
11      Health Technology  6.66   6.70  2.95 10.70 -30.8 28.9
12    Industrial Services  6.58   5.70  4.20  9.00   3.3 10.6
13    Non-Energy Minerals  7.10   5.20  1.05 11.70  -2.2 21.2
14     Process Industries  7.01   6.25  5.32  7.43  -5.0 24.8
15 Producer Manufacturing  9.39   8.85  5.55 10.95  -3.1 25.1
16           Retail Trade  8.84   8.95  5.38 13.65 -23.2 24.8
17    Technology Services  9.24   8.65  5.15 13.75 -35.4 41.5
18         Transportation  8.77   7.90  3.95 10.70   1.1 26.2
19              Utilities  2.33   2.30  2.00  3.20  -6.5  5.6
```

2. Box Plot for ROA by Sector

```r
library(ggplot2)

ggplot(sp500, aes(Sector, ROA)) + geom_boxplot() +
  theme(axis.text.x = element_text(angle = 45, vjust = 1, hjust = 1))
```



Live Case: S&P500 (2b of 3)

Aug 06, 2023 -=- This chapter is being heavily edited

ISSUE: Analysis of a particular SECTOR We have chosen to deeply analyze the HEALTH TECHNOLOGY Sector

SECTOR LEVEL ANALYSIS begins here

Filter the data by sector Health Services, and display the number of stocks in the sector

```
ts <- sp500 %>%
        filter(Sector=='Health Services')

nrow(ts)
```

```
[1] 12
```

There are 12 number of of stocks in the sector Health Services

Select the Specific Coulumns from the filtered dataframe `ts` (Health Services)

```
ts2 <- ts %>%
        select(Date, Stock, StockName,Sector,  Industry, MarketCap, Price,Low52Wk, High52W
               ROE, ROA,ROIC,GrossMargin, GrossMargin,
               NetMargin, Rating)

colnames(ts2)
```

```
 [1] "Date"        "Stock"       "StockName"   "Sector"      "Industry"
 [6] "MarketCap"   "Price"       "Low52Wk"     "High52Wk"    "ROE"
[11] "ROA"         "ROIC"        "GrossMargin" "NetMargin"   "Rating"
```

Arrange the Dataframe by ROE

```
ts3 <- ts2 %>% arrange(desc(ROE))
```

Top 10 Shares in Sector Health Services Based on ROE

```
head(ts3,10)
```

```
# A tibble: 10 x 15
   Date   Stock StockName Sector Industry MarketCap Price Low52Wk High52Wk   ROE
   <chr>  <chr> <chr>     <fct>  <chr>        <dbl> <dbl>   <dbl>    <dbl> <dbl>
 1 9/30/~ DVA   DaVita I~ Healt~ Medical~   8.63e 9  94.6    65.3    117    56
 2 9/30/~ MOH   Molina H~ Healt~ Managed~   1.91e10 328.    256.     374    28.4
 3 9/30/~ UNH   UnitedHe~ Healt~ Managed~   4.67e11 504.    446.     558.   27.2
 4 9/30/~ HUM   Humana I~ Healt~ Managed~   6.03e10 486.    423.     571.   20.9
 5 9/30/~ IQV   IQVIA Ho~ Healt~ Service~   3.60e10 197.    166.     242.   19.7
 6 9/30/~ ELV   Elevance~ Healt~ Managed~   1.03e11 435.    412      550.   17.3
 7 9/30/~ CI    The Cign~ Healt~ Managed~   8.47e10 286.    240.     340.   14.6
 8 9/30/~ DGX   Quest Di~ Healt~ Service~   1.37e10 122.    122.     158.   12.5
 9 9/30/~ UHS   Universa~ Healt~ Hospita~   7.81e 9 126.     82.5    159.   11.6
10 9/30/~ CNC   Centene ~ Healt~ Managed~   3.73e10  68.9    60.8     87.8  10.4
# i 5 more variables: ROA <dbl>, ROIC <dbl>, GrossMargin <dbl>,
#   NetMargin <dbl>, Rating <fct>
```

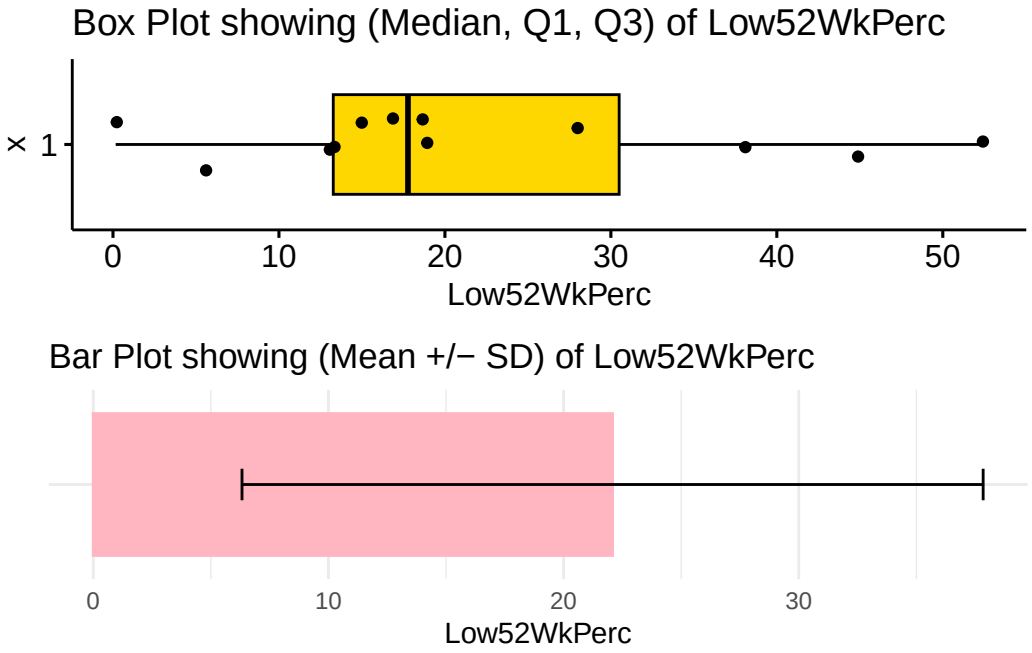
Mutate a data column called (Low52WkPerc), then show top 10 ROE stocks

```
ts4 <- ts3 %>% mutate(Low52WkPerc = round((Price - Low52Wk)*100 / Low52Wk,2))
head(ts4[,c(1:3,10,16)],10)
```

```
# A tibble: 10 x 5
   Date      Stock StockName                          ROE Low52WkPerc
   <chr>     <chr> <chr>                            <dbl>       <dbl>
 1 9/30/2023 DVA   DaVita Inc.                       56          44.9
 2 9/30/2023 MOH   Molina Healthcare Inc             28.4        28.0
 3 9/30/2023 UNH   UnitedHealth Group Incorporated   27.2        13.1
 4 9/30/2023 HUM   Humana Inc.                       20.9        14.9
 5 9/30/2023 IQV   IQVIA Holdings, Inc.              19.7        18.7
```

```
 6 9/30/2023 ELV   Elevance Health, Inc.            17.3        5.68
 7 9/30/2023 CI    The Cigna Group                  14.6       19.0
 8 9/30/2023 DGX   Quest Diagnostics Incorporated   12.5        0.16
 9 9/30/2023 UHS   Universal Health Services, Inc.  11.6       52.4
10 9/30/2023 CNC   Centene Corporation              10.4       13.3
```

Low52WkPerc for all the Health Sector Stocks, as shown below



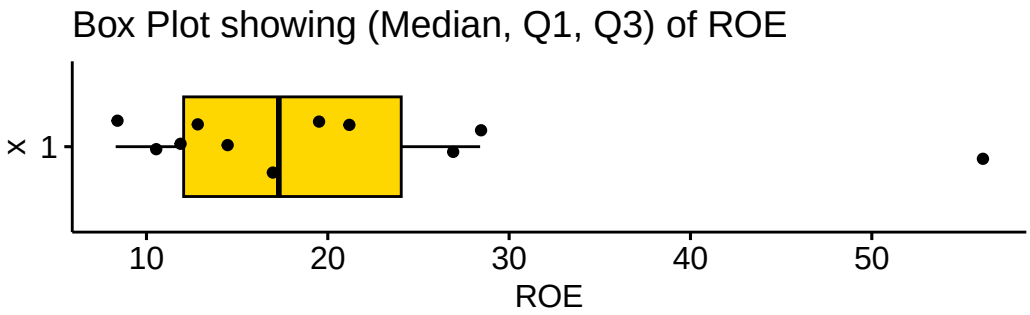
Summary Statistics of ROE

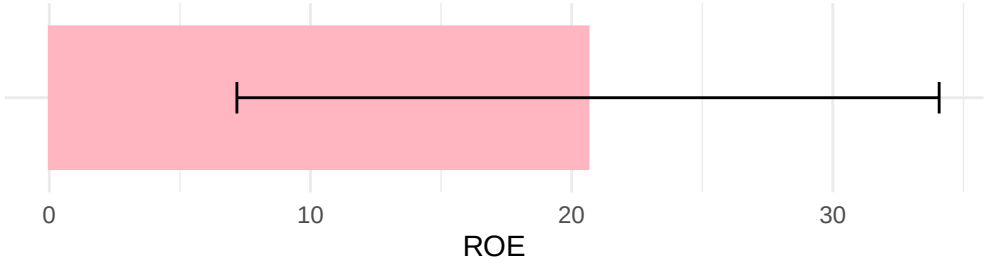
```r
ts3 <- na.omit(ts3)

ROESum <- ts3 %>%
  summarise(
    Mean = mean(ROE),
    Median= sd(ROE),
    Median= median(ROE),
    Q1 = quantile(ROE, probs = 0.25, na.rm = TRUE),
    Q3 = quantile(ROE, probs = 0.75, na.rm = TRUE),
    Min = min(ROE),
    max = max(ROE)
  )
```

```
ROESum <- round(ROESum,2)
ROESum
```

```
# A tibble: 1 x 6
   Mean Median    Q1    Q3   Min   max
  <dbl>  <dbl> <dbl> <dbl> <dbl> <dbl>
1  21.6   12.5  11.6  19.7   8.3    56
```

- ROE for all the Stocks in Health Sector, as shown below*





Top 10 Shares in Health Sector with highest ROE

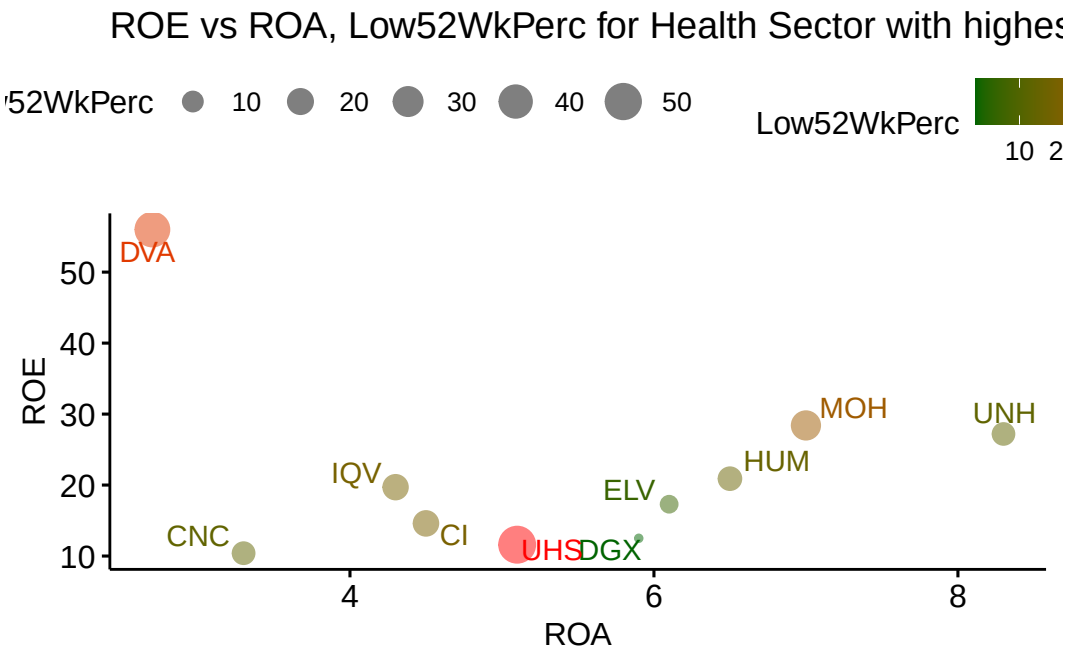
```
ts4 %>%
  select(Stock, Price, Low52Wk, Low52WkPerc, ROA, ROE) %>%
  arrange(desc(ROE))%>%
  slice(1:10) %>%
  kable("html", caption = "Top 10 Shares in Health Sector with highest ROE") %>%
  kable_styling()
```

Table 0.5: Top 10 Shares in Health Sector with highest ROE

Stock	Price	Low52Wk	Low52WkPerc	ROA	ROE
DVA	94.6	65.3	44.87	2.7	56.0
MOH	327.9	256.2	27.99	7.0	28.4
UNH	504.2	445.7	13.13	8.3	27.2
HUM	486.5	423.3	14.93	6.5	20.9
IQV	196.8	165.8	18.70	4.3	19.7
ELV	435.4	412.0	5.68	6.1	17.3
CI	286.1	240.5	18.96	4.5	14.6
DGX	121.9	121.7	0.16	5.9	12.5
UHS	125.7	82.5	52.36	5.1	11.6
CNC	68.9	60.8	13.32	3.3	10.4

ROE versus ROA and colored by Price rel. to 52 Week Low

```
top10 <-
  ts4 %>%
  select(Stock, Price, Low52Wk, Low52WkPerc, ROA, ROE) %>%
  arrange(desc(ROE))%>%
  slice(1:10)

top10$name <- top10$Stock

ggscatter(top10,
          x = "ROA",
          y = "ROE",
          size = "Low52WkPerc",
          color = "Low52WkPerc",
          alpha = 0.5,
          label = "name",
          repel = TRUE,
          title = "ROE vs ROA, Low52WkPerc for Health Sector with highest ROE") +
  gradient_color(c("darkgreen",  "red"))
```



Summary Statistics of All key variables in Sector Health Services

```
ts3 <- na.omit(ts3)

ROESum <- ts3 %>%
  summarise(
    Mean = mean(ROE),
    Median= sd(ROE),
    Median= median(ROE),
    Q1 = quantile(ROE, probs = 0.25, na.rm = TRUE),
    Q3 = quantile(ROE, probs = 0.75, na.rm = TRUE),
    Min = min(ROE),
    max = max(ROE)
  )

ROESum <- round(ROESum,2)


ROASum <- ts3 %>%
  summarise(
    Mean = mean(ROA),
    Median= sd(ROA),
```

```
    Median= median(ROA),
    Q1 = quantile(ROA, probs = 0.25, na.rm = TRUE),
    Q3 = quantile(ROA, probs = 0.75, na.rm = TRUE),
    Min = min(ROA),
    max = max(ROA)
  )

ROASum <- round(ROASum,2)

ROICSum <- ts3 %>%
  summarise(
    Mean = mean(ROIC),
    Median= sd(ROIC),
    Median= median(ROIC),
    Q1 = quantile(ROIC, probs = 0.25, na.rm = TRUE),
    Q3 = quantile(ROIC, probs = 0.75, na.rm = TRUE),
    Min = min(ROIC),
    max = max(ROIC)
  )

ROICSum <- round(ROICSum,2)

GrossMarginSum <- ts3 %>%
  summarise(
    Mean = mean(GrossMargin),
    Median= sd(GrossMargin),
    Median= median(GrossMargin),
    Q1 = quantile(GrossMargin, probs = 0.25, na.rm = TRUE),
    Q3 = quantile(GrossMargin, probs = 0.75, na.rm = TRUE),
    Min = min(GrossMargin),
    max = max(GrossMargin)
  )

GrossMarginSum <- round(GrossMarginSum,2)

NetMarginSum <- ts3 %>%
  summarise(
    Mean = mean(NetMargin),
    Median= sd(NetMargin),
    Median= median(NetMargin),
    Q1 = quantile(NetMargin, probs = 0.25, na.rm = TRUE),
```

```
    Q3 = quantile(NetMargin, probs = 0.75, na.rm = TRUE),
    Min = min(NetMargin),
    max = max(NetMargin)
  )

NetMarginSum <- round(NetMarginSum,2)

Metrics <- c("ROE","ROA","ROIC","GrossMargin","NetMargin")

ftab <- rbind(ROESum, ROASum, ROICSum, GrossMarginSum, NetMarginSum)
ftab <- cbind(Metrics, ftab)
ftab
```

```
      Metrics  Mean Median   Q1   Q3 Min  max
1         ROE 21.62   12.5 11.6 19.7 8.3 56.0
2         ROA  4.44    4.3  4.2  5.1 2.7  5.9
3        ROIC  5.70    6.0  5.1  6.3 3.7  7.4
4 GrossMargin 23.26   25.5 23.0 27.1 7.9 32.8
5   NetMargin  6.08    5.7  5.0  7.5 3.9  8.3
```

Summary Statistics of ROE by each Sector of S&P500

```
SectorROE <- sp500 %>%
  group_by(Sector) %>%
  summarise(
    Mean = mean(na.omit(ROE)),
    Median= sd(na.omit(ROE)),
    Median= median(na.omit(ROE)),
    Q1 = quantile(na.omit(ROE), probs = 0.25, na.rm = TRUE),
    Q3 = quantile(na.omit(ROE), probs = 0.75, na.rm = TRUE),
    Min = min(na.omit(ROE)),
    max = max(na.omit(ROE))
  )

cbind(Sector = SectorROE$Sector, round(SectorROE[,2:7],2))
```

```
                  Sector   Mean Median    Q1    Q3     Min    max
1     Commercial Services  37.98  26.40 16.40 43.60     3.5  175.2
2          Communications   8.10   9.10  0.55 16.15    -8.0   23.2
```

```
3       Consumer Durables  12.23  16.65  6.85 25.38   -51.4   45.2
4   Consumer Non-Durables 129.53  19.60  6.40 33.90   -11.5 2878.8
5       Consumer Services  33.02  11.40  1.55 44.95  -185.6  359.9
6   Distribution Services  81.10  34.20 22.15 56.45     5.1  371.2
7   Electronic Technology  31.51  18.75  8.10 36.80   -14.8  160.1
8         Energy Minerals  43.12  26.95 23.78 41.45    18.0  230.2
9                 Finance  22.13  11.00  7.82 16.67   -14.7  714.3
10        Health Services  20.63  17.30 12.05 24.05     8.3   56.0
11      Health Technology  19.87  13.10  6.80 22.73   -49.3  173.5
12    Industrial Services  21.04  22.60 10.70 31.10     7.7   36.5
13    Non-Energy Minerals  13.84  13.50  3.40 21.80    -3.8   36.8
14     Process Industries  25.72  18.60 15.35 24.62   -13.2  125.5
15 Producer Manufacturing  25.42  20.20 13.02 30.00   -13.6   95.9
16           Retail Trade  74.34  28.70 14.47 44.00 -1224.5 2065.3
17    Technology Services  33.17  18.00 10.70 31.82   -70.6  416.6
18         Transportation  36.39  33.50 20.85 49.08     4.1  104.4
19              Utilities   8.12   8.70  7.65 10.60   -47.6   35.5
```

ANALYSIS OF HEALTH SERVICES SECTOR

1. Market Cap of all companies in Sector Health Services

```r
library(janitor)
library(kableExtra)
# Market Cap by Stock
MCap <- ts3 %>%
  group_by(Stock) %>%
  summarise(
    MarketCapCr = sum(na.omit(MarketCap)/10000000))

# Sp500 Market Cap

SP500MarketCap <- sum(ts3$MarketCap/10000000)

# calculating % market cap
PercentMarketCap <- round(MCap$MarketCapCr*100/SP500MarketCap,2)
MCapTab <- cbind(MCap,PercentMarketCap)

# sorting by PercentMarketCap
MCapTab <- MCapTab %>% arrange(desc(PercentMarketCap))
```

```
MCapTab <- MCapTab %>%
  adorn_totals("row")

MCapTab <- knitr::kable(MCapTab, "html") %>% kable_styling()
MCapTab
```

Stock	MarketCapCr	PercentMarketCap
IQV	3602.9303	42.91
LH	1781.3030	21.22
DGX	1367.6957	16.29
DVA	863.0589	10.28
UHS	781.2901	9.31
Total	8396.2781	100.01

2. Shares which are most attractively priced in Sector Health Services

```
AttrShares <- ts4 %>% arrange(Low52WkPerc)
AttrShares <- AttrShares[, c(2:4,7,8,10,11,16)]

AttrShares <- knitr::kable(AttrShares, "html") %>% kable_styling()
AttrShares
```

Stock	StockName	Sector	Price	Low52Wk	ROE	ROA	Lo
DGX	Quest Diagnostics Incorporated	Health Services	121.9	121.7	12.5	5.9	
ELV	Elevance Health, Inc.	Health Services	435.4	412.0	17.3	6.1	
UNH	UnitedHealth Group Incorporated	Health Services	504.2	445.7	27.2	8.3	
CNC	Centene Corporation	Health Services	68.9	60.8	10.4	3.3	
HUM	Humana Inc.	Health Services	486.5	423.3	20.9	6.5	
LH	Laboratory Corporation of America Holdings	Health Services	201.1	172.1	8.3	4.2	
IQV	IQVIA Holdings, Inc.	Health Services	196.8	165.8	19.7	4.3	
CI	The Cigna Group	Health Services	286.1	240.5	14.6	4.5	
MOH	Molina Healthcare Inc	Health Services	327.9	256.2	28.4	7.0	
HCA	HCA Healthcare, Inc.	Health Services	246.1	178.3	NA	11.0	
DVA	DaVita Inc.	Health Services	94.6	65.3	56.0	2.7	
UHS	Universal Health Services, Inc.	Health Services	125.7	82.5	11.6	5.1	

PROFITABILITY OF HEALTH SERVICES SECTOR

1. Shares have highest ROE within Sector Technology Services

```
AttrShares <- ts4 %>% arrange(desc(ROE))
AttrShares <- AttrShares[, c(2:4,7,8,10,11,16)]

AttrShares <- knitr::kable(AttrShares, "html") %>% kable_styling()
AttrShares
```

Stock	StockName	Sector	Price	Low52Wk	ROE	ROA	L
DVA	DaVita Inc.	Health Services	94.6	65.3	56.0	2.7	
MOH	Molina Healthcare Inc	Health Services	327.9	256.2	28.4	7.0	
UNH	UnitedHealth Group Incorporated	Health Services	504.2	445.7	27.2	8.3	
HUM	Humana Inc.	Health Services	486.5	423.3	20.9	6.5	
IQV	IQVIA Holdings, Inc.	Health Services	196.8	165.8	19.7	4.3	
ELV	Elevance Health, Inc.	Health Services	435.4	412.0	17.3	6.1	
CI	The Cigna Group	Health Services	286.1	240.5	14.6	4.5	
DGX	Quest Diagnostics Incorporated	Health Services	121.9	121.7	12.5	5.9	
UHS	Universal Health Services, Inc.	Health Services	125.7	82.5	11.6	5.1	
CNC	Centene Corporation	Health Services	68.9	60.8	10.4	3.3	
LH	Laboratory Corporation of America Holdings	Health Services	201.1	172.1	8.3	4.2	
HCA	HCA Healthcare, Inc.	Health Services	246.1	178.3	NA	11.0	

2. Shares have highest ROA within Sector Health Services

```
AttrShares <- ts4 %>% arrange(desc(ROA))
AttrShares <- AttrShares[, c(2:4,7,8,10,11,16)]

AttrShares <- knitr::kable(AttrShares, "html") %>% kable_styling()
AttrShares
```

Stock	StockName	Sector	Price	Low52Wk	ROE	ROA	L
HCA	HCA Healthcare, Inc.	Health Services	246.1	178.3	NA	11.0	
UNH	UnitedHealth Group Incorporated	Health Services	504.2	445.7	27.2	8.3	
MOH	Molina Healthcare Inc	Health Services	327.9	256.2	28.4	7.0	
HUM	Humana Inc.	Health Services	486.5	423.3	20.9	6.5	
ELV	Elevance Health, Inc.	Health Services	435.4	412.0	17.3	6.1	
DGX	Quest Diagnostics Incorporated	Health Services	121.9	121.7	12.5	5.9	
UHS	Universal Health Services, Inc.	Health Services	125.7	82.5	11.6	5.1	

Stock	StockName	Sector	Price	Low52Wk	ROE	ROA	Lo
CI	The Cigna Group	Health Services	286.1	240.5	14.6	4.5	
IQV	IQVIA Holdings, Inc.	Health Services	196.8	165.8	19.7	4.3	
LH	Laboratory Corporation of America Holdings	Health Services	201.1	172.1	8.3	4.2	
CNC	Centene Corporation	Health Services	68.9	60.8	10.4	3.3	
DVA	DaVita Inc.	Health Services	94.6	65.3	56.0	2.7	

3. Shares have highest NetMargin within Sector Health Services

```
AttrShares <- ts4 %>% arrange(desc(NetMargin))
AttrShares <- AttrShares[, c(2:4,7,8,10,11,14,16)]

AttrShares <- knitr::kable(AttrShares, "html") %>% kable_styling()
AttrShares
```

Stock	StockName	Sector	Price	Low52Wk	ROE	ROA	N
HCA	HCA Healthcare, Inc.	Health Services	246.1	178.3	NA	11.0	
DGX	Quest Diagnostics Incorporated	Health Services	121.9	121.7	12.5	5.9	
IQV	IQVIA Holdings, Inc.	Health Services	196.8	165.8	19.7	4.3	
UNH	UnitedHealth Group Incorporated	Health Services	504.2	445.7	27.2	8.3	
LH	Laboratory Corporation of America Holdings	Health Services	201.1	172.1	8.3	4.2	
UHS	Universal Health Services, Inc.	Health Services	125.7	82.5	11.6	5.1	
DVA	DaVita Inc.	Health Services	94.6	65.3	56.0	2.7	
ELV	Elevance Health, Inc.	Health Services	435.4	412.0	17.3	6.1	
CI	The Cigna Group	Health Services	286.1	240.5	14.6	4.5	
HUM	Humana Inc.	Health Services	486.5	423.3	20.9	6.5	
MOH	Molina Healthcare Inc	Health Services	327.9	256.2	28.4	7.0	
CNC	Centene Corporation	Health Services	68.9	60.8	10.4	3.3	