Case Analysis of SP500

We will investigate data related to the S&P500 stocks.

This is the URL of the data we will use: https://docs.google.com/spreadsheets/d/11ahk9uWxBkDqrhNm7qYmiT

1) Reading S&P500 data from a Google Sheet into a tibble

The Google Sheet ID is: `11ahk9uWxBkDqrhNm7qYmiTwrlSC53N1zvXYfv7ttOCM`. We can use the function `gsheet2tbl` in package `gsheet` to read the Google Sheet into a dataframe, as demonstrated in the following code.

```
# Read S&P500 stock data present in a Google Sheet.
library(gsheet)

prefix <- "https://docs.google.com/spreadsheets/d/"
sheetID <- "11ahk9uWxBkDqrhNm7qYmiTwrlSC53N1zvXYfv7ttOCM"
# Form the URL to connect to
url500 <- paste(prefix, sheetID)
# Read the Google Sheet located at the URL into a tibble called sp500
sp500 <- gsheet2tbl(url500)
```

```
No encoding supplied: defaulting to UTF-8.
```

1) Reviewing the data

We run the str() function to better understand the data.

```
str(sp500)
```

```
Classes 'tbl_df', 'tbl' and 'data.frame':   503 obs. of  36 variables:
 $ Date                                  : chr  "7/13/2023" "7/13/2023" "7/13/2023" "7/13/20
 $ Stock                                 : chr  "A" "AAL" "AAP" "AAPL" ...
 $ Description                           : chr  "Agilent Technologies, Inc." "American Airli
 $ Sector                               : chr  "Health Technology" "Transportation" "Retail
 $ Industry                             : chr  "Medical Specialties" "Airlines" "Specialty
 $ Market.Capitalization                : num  3.54e+10 1.20e+10 4.19e+09 2.99e+12 2.36e+1
 $ Price                                : num  120 18.4 70.5 190.2 133.6 ...
 $ X52.Week.Low                         : num  113.3 11.7 63.6 124.2 131 ...
 $ X52.Week.High                        : num  160 19.1 212 194 168 194 116 78.3 328 519 .
 $ Return.on.Equity..TTM.               : num  24.8 NA 14.6 146 51.1 389 16 14.8 30.7 33.7
 $ Return.on.Assets..TTM.               : num  12.7 2.64 3.35 27.6 5.43 2.79 7.82 4.98 14.
 $ Return.on.Invested.Capital..TTM.     : num  16.51 5.44 6.17 57.18 9.9 ...
 $ Gross.Margin..TTM.                   : num  54.1 21.7 43.8 43.2 72.2 ...
 $ Operating.Margin..TTM.               : num  23.78 7.3 5.63 29.16 41.07 ...
 $ Net.Margin..TTM.                     : num  19.19 3.39 3.61 24.49 13.3 ...
 $ Price.to.Earnings.Ratio..TTM.        : num  26.45 7.23 10.48 32.32 31.46 ...
 $ Price.to.Book..FY.                   : num  6.68 NA 1.56 59.85 13.7 ...
 $ Enterprise.Value.EBITDA..TTM.        : num  18.8 7.08 8.85 24.5 9.24 12.8 19.1 NA 17.4
 $ EBITDA..TTM.                         : num  1.97e+09 6.02e+09 9.21e+08 1.24e+11 3.18e+1
 $ EPS.Diluted..TTM.                    : num  4.54 2.54 6.72 5.89 4.25 ...
 $ EBITDA..TTM.YoY.Growth.              : num  10.52 NA -16 -5.36 10.6 ...
 $ EBITDA..Quarterly.YoY.Growth.        : num  8.2 NA -39.01 -4.58 11.68 ...
 $ EPS.Diluted..TTM.YoY.Growth.         : num  9.17 NA -25.21 -4.33 -39.11 ...
 $ EPS.Diluted..Quarterly.YoY.Growth.   : num  11.69944 NA -68.3683 -0.00656 -94.8904 ...
 $ Price.to.Free.Cash.Flow..TTM.        : num  30.13 7.03 NA 30.92 10.09 ...
 $ Free.Cash.Flow..TTM.YoY.Growth.      : num  11.81 NA -100.23 -7.85 6.68 ...
 $ Free.Cash.Flow..Quarterly.YoY.Growth.  : num  55.7078 648.1481 -176.135 -0.0312 -15.3392
 $ Debt.to.Equity.Ratio..MRQ.           : num  0.473 NA 1.582 1.763 4.678 ...
 $ Current.Ratio..MRQ.                  : num  2.37 0.718 1.244 0.94 0.96 ...
 $ Quick.Ratio..MRQ.                    : num  1.708 0.624 0.238 0.878 0.821 ...
 $ Dividend.Yield.Forward               : num  0.75 NA 1.413 0.506 4.386 ...
 $ Dividends.per.share..Annual.YoY.Growth.: num  8.25 NA 84.62 5.88 7.53 ...
 $ Price.to.Sales..FY.                  : num  5.257 0.246 0.383 7.875 4.091 ...
 $ Revenue..TTM.YoY.Growth.             : num  7.8597 50.2948 1.4153 -0.2544 0.0282 ...
 $ Revenue..Quarterly.YoY.Growth.       : num  6.85 36.97 1.29 -2.51 -9.7 ...
 $ Technical.Rating                     : chr  "Neutral" "Sell" "Buy" "Sell" ...
```

The `str(sp500)` output provides valuable insights into the structure and data types of the columns in the `sp500` tibble. Let's delve into the details:

The output reveals that `sp500` is a tibble with dimensions [503 × 36]. This means it consists of 503 rows, each representing a specific S&P500 stock, and 36 columns containing information about each stock.

Here's a breakdown of the information associated with each column:

- The columns labeled `Date`, `Stock`, `Description`, `Sector`, and `Industry` are character columns. They respectively represent the date, stock ticker symbol, description, sector, and industry of each S&P500 stock.
- Columns such as `Market.Capitalization`, `Price`, `X52.Week.Low`, `X52.Week.High`, and other numeric columns contain diverse financial metrics and stock prices related to the S&P500 stocks.
- The column labeled `Technical.Rating` is a character column that assigns a technical rating to each stock.

By examining the `str(sp500)` output, you gain a comprehensive understanding of the data types and column names present in the `sp500` tibble, enabling you to grasp the structure of the dataset effectively.

```r
library(dplyr)
```

```

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```r
# Define a mapping of new column names
new_names <- c(
  "Date", "Stock", "Desc", "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", "TechRating"
)
```

```r
# Rename the columns using the new_names vector
sp500_renamed <- sp500 %>%
  rename_with(~ new_names, everything())

# Print the updated column names
names(sp500_renamed)
```

```
 [1] "Date"                  "Stock"                 "Desc"
 [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"          "TechRating"
```

```r
library(tibble)

new_names <- c(
  "Date", "Stock", "Desc", "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", "TechRating"
)

descriptions <- c(
  "Date of the observation",
  "Stock ticker symbol",
  "Description of the stock",
  "Sector the stock belongs to",
  "Industry the stock belongs to",
  "Market capitalization of the company",
  "Stock price",
  "52-week low price",
  "52-week high price",
  "Return on Equity",
  "Return on Assets",
  "Return on Invested Capital",
  "Gross Margin",
  "Operating Margin",
  "Net Margin",
  "Price-to-Earnings Ratio",
  "Price-to-Book Ratio",
  "Enterprise Value to EBITDA",
  "EBITDA",
  "Earnings per Share",
  "EBITDA Year-over-Year Growth",
  "EBITDA Quarterly Year-over-Year Growth",
  "EPS Year-over-Year Growth",
  "EPS Quarterly Year-over-Year Growth",
  "Price-to-Free Cash Flow",
  "Free Cash Flow",
  "Free Cash Flow Quarterly Year-over-Year Growth",
  "Debt-to-Equity Ratio",
  "Current Ratio",
  "Quick Ratio",
  "Dividend Yield",
  "Annual Dividends per Share Year-over-Year Growth",
  "Price-to-Sales Ratio",
  "Revenue Year-over-Year Growth",
  "Revenue Quarterly Year-over-Year Growth",
  "Technical Rating"
)

column_info <- tibble(NewName = new_names, Description = descriptions)

print(column_info)
```

```
# A tibble: 36 x 2
   NewName   Description
   <chr>     <chr>
 1 Date      Date of the observation
 2 Stock     Stock ticker symbol
 3 Desc      Description of the stock
 4 Sector    Sector the stock belongs to
 5 Industry  Industry the stock belongs to
 6 MarketCap Market capitalization of the company
 7 Price     Stock price
 8 Low52Wk   52-week low price
 9 High52Wk  52-week high price
10 ROE       Return on Equity
# i 26 more rows
```