BA 64060 - Assignment 1

Disney Maxwell

2025-09-06

1. Dataset: Forbes – The Global 2000 Companies – 2025

This dataset is from kaggle.com and contains financial data of the world's top 2000 companies.
Dataset link: https://www.kaggle.com/datasets/ellimaaac/forbes-the-global-2000-companies-2025.

2. Import Dataset into R

Import financial data into R.

```
library(readr)
forbes_2000_companies <- read_csv("Forbes_2000_Companies_2025.csv")
```

```
## Rows: 2000 Columns: 8
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (3): Company, Headquarters, Industry
## dbl (5): Rank, Sales ($B), Profit ($B), Assets ($B), Market Value ($B)
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
View(forbes_2000_companies)
```

3.A. Descriptive Statistics - Quantitative Variables

i. Average Sales of 2000 Companies

```
sales_df <- forbes_2000_companies["Sales ($B)"]
mean_sales_value <- colMeans(sales_df)
print(paste0("The Mean sales value of the 2000 companies = ", mean_sales_value))
```

```
## [1] "The Mean sales value of the 2000 companies = 26.44163"
```

ii. Range Profit of 2000 Companies

```
profit_vector <- forbes_2000_companies[["Profit ($B)"]]
range_profit_value <- max(profit_vector, na.rm = TRUE) - min(profit_vector, na.rm = TRUE)
print(paste0("The Range value for profit of the 2000 companies = ", range_profit_value))
```

```
## [1] "The Range value for profit of the 2000 companies = 132.65"
```

3.B. Descriptive Statistics - Categorical Variables

i. Number of Company Headquarters in different Countries

```
headqrtr_freq <- table(forbes_2000_companies["Headquarters"])
print(paste0("Number of Headquarters across the different Countries:"))
```

```
## [1] "Number of Headquarters across the different Countries:"
```

```
print(headqrtr_freq)
```

```
## Headquarters
##             Argentina             Australia               Austria
##                     3                    32                     9
##               Bahrain               Belgium               Bermuda
##                     1                     7                     6
##                Brazil                Canada        Cayman Islands
##                    27                    60                     2
##                 Chile                 China              Colombia
##                     6                   275                     8
##                Cyprus        Czech Republic               Denmark
##                     1                     1                    13
##                 Egypt               Finland                France
##                     1                    10                    47
##               Germany                Greece             Hong Kong
##                    49                     6                    42
##               Hungary                 India             Indonesia
##                     3                    70                    12
##               Ireland                Israel                 Italy
##                    25                    13                    30
##                 Japan                Jordan            Kazakhstan
##                   180                     1                     2
##                Kuwait               Lebanon            Luxembourg
##                     2                     1                     4
##              Malaysia                Mexico               Morocco
##                     9                    12                     2
##           Netherlands                Norway                  Oman
##                    23                     8                     1
##                Panama                  Peru           Philippines
##                     1                     1                     7
##                Poland              Portugal                 Qatar
##                     8                     4                     6
##               Romania          Saudi Arabia             Singapore
##                     1                    19                    11
##              Slovenia          South Africa           South Korea
##                     1                    14                    62
##                 Spain                Sweden           Switzerland
##                    19                    27                    45
##                Taiwan              Thailand                Turkey
##                    41                    16                    10
## United Arab Emirates        United Kingdom         United States
##                    15                    68                   612
##               Vietnam
##                     8
```

ii. Country with the most number of Headquarters

```
most_freq_headqrtr <- names(which.max(headqrtr_freq))
print(paste0("Country with the most number of headquarters: ", most_freq_headqrtr))
```

```
## [1] "Country with the most number of headquarters: United States"
```

4. Transformations - Convert Sales from Billion to Million dollar values

```
sales_millions_df <- sales_df * 1000
names(sales_millions_df) <- c("Sales ($M)")
first_10_companies_sales = head(sales_millions_df, 10)
print(paste0("First 10 Company Sales in million dollars"))
```

```
## [1] "First 10 Company Sales in million dollars"
```

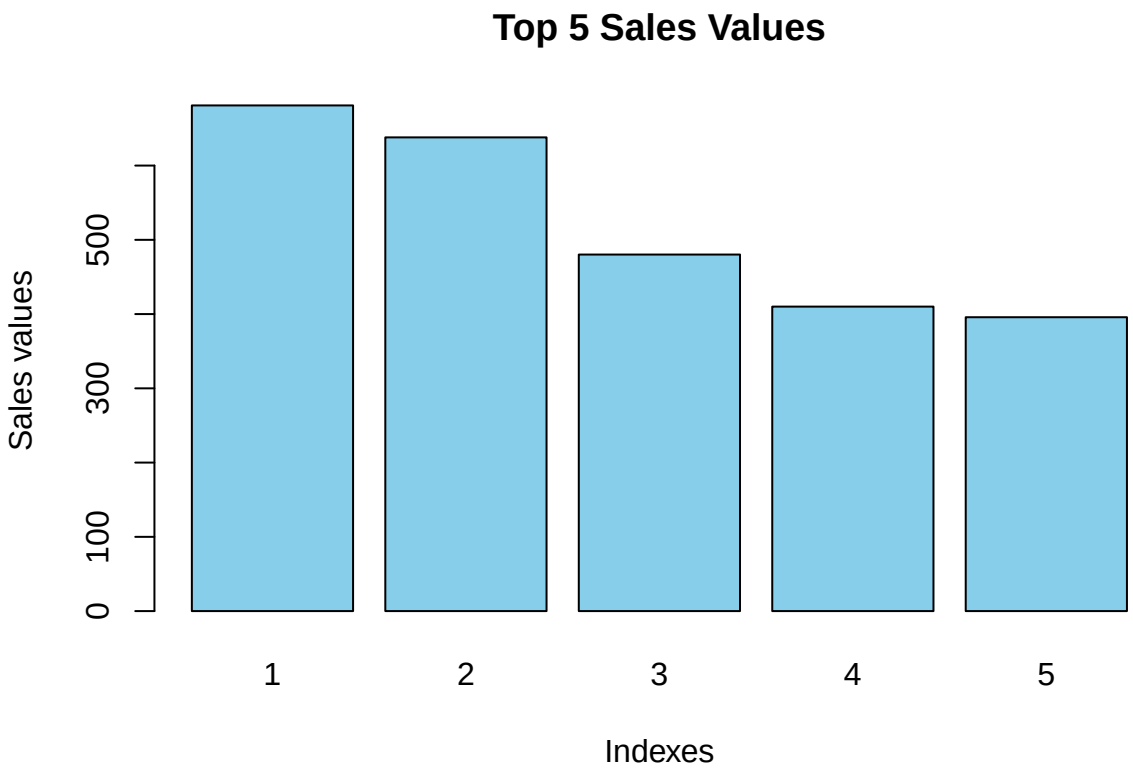
```
print(first_10_companies_sales)
```

```
##    Sales ($M)
## 1      285110
## 2      371430
## 3      221960
## 4      480150
## 5      637960
## 6      196530
## 7      196710
## 8      198020
## 9      359310
## 10     261800
```

5.A. Plot - Quantitative Variable - Top 5 Sales Values

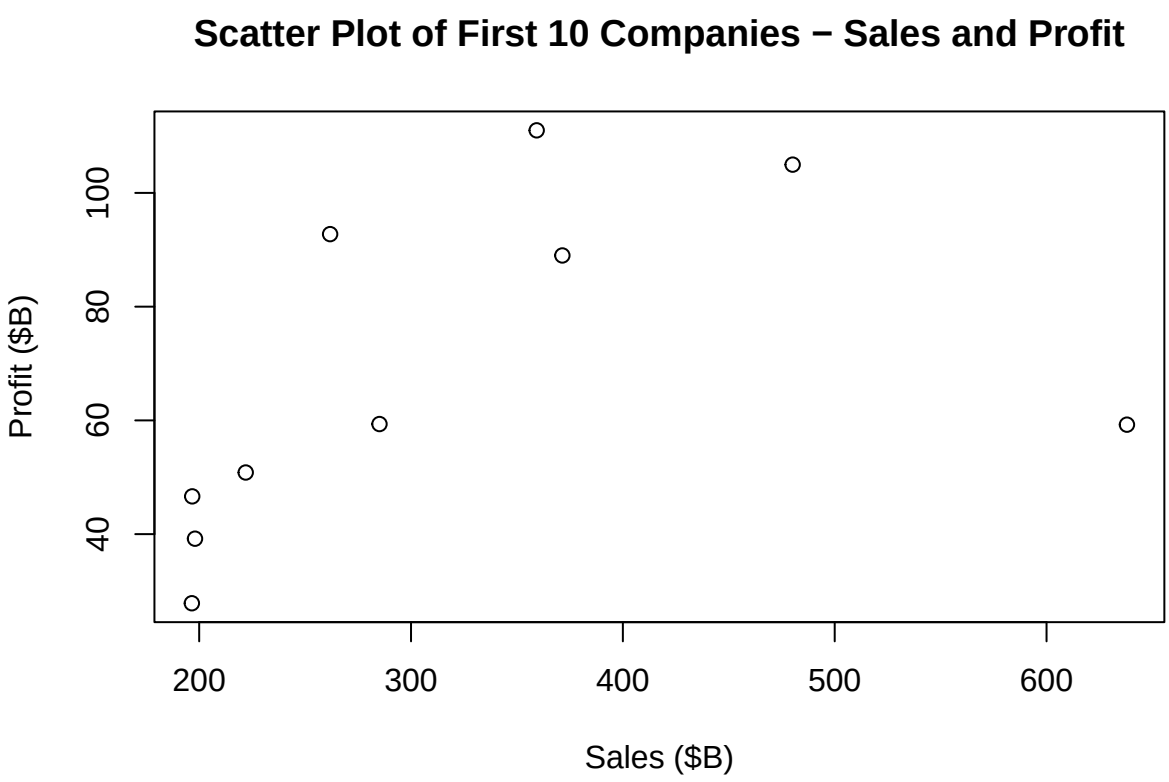
```
sales_vector <- forbes_2000_companies[["Sales ($B)"]]
top_5_sales_vector <- sort(sales_vector, decreasing = TRUE)[1:5]
bar_x_labels <- c("1", "2", "3", "4", "5")

# Plot the vector as a bar chart
barplot(top_5_sales_vector,
        names.arg = bar_x_labels, # Assign labels to bars
        main = "Top 5 Sales Values", # Set the main title of the plot
        xlab = "Indexes", # Set the x-axis label
        ylab = "Sales values", # Set the y-axis label
        col = "skyblue" # Set the color of the bars
        )
```



5.B. Scatter Plot - First 10 Companies - Sales and Profit

```
first_10_companies <- forbes_2000_companies[1:10, c("Sales ($B)", "Profit ($B)")]
plot(first_10_companies$`Sales ($B)`, first_10_companies$`Profit ($B)`,
     xlab = "Sales ($B)", ylab = "Profit ($B)",
     main = "Scatter Plot of First 10 Companies - Sales and Profit")
```


Scatter Plot of First 10 Companies − Sales and Profit
Profit ($B)
40
60
80
100
200
300
400
500
600
Sales ($B)