Question 1

January 28, 2021

```
[1]: library(readxl)
     data <- read_excel("Forbes-ten-companies.xlsx")
```
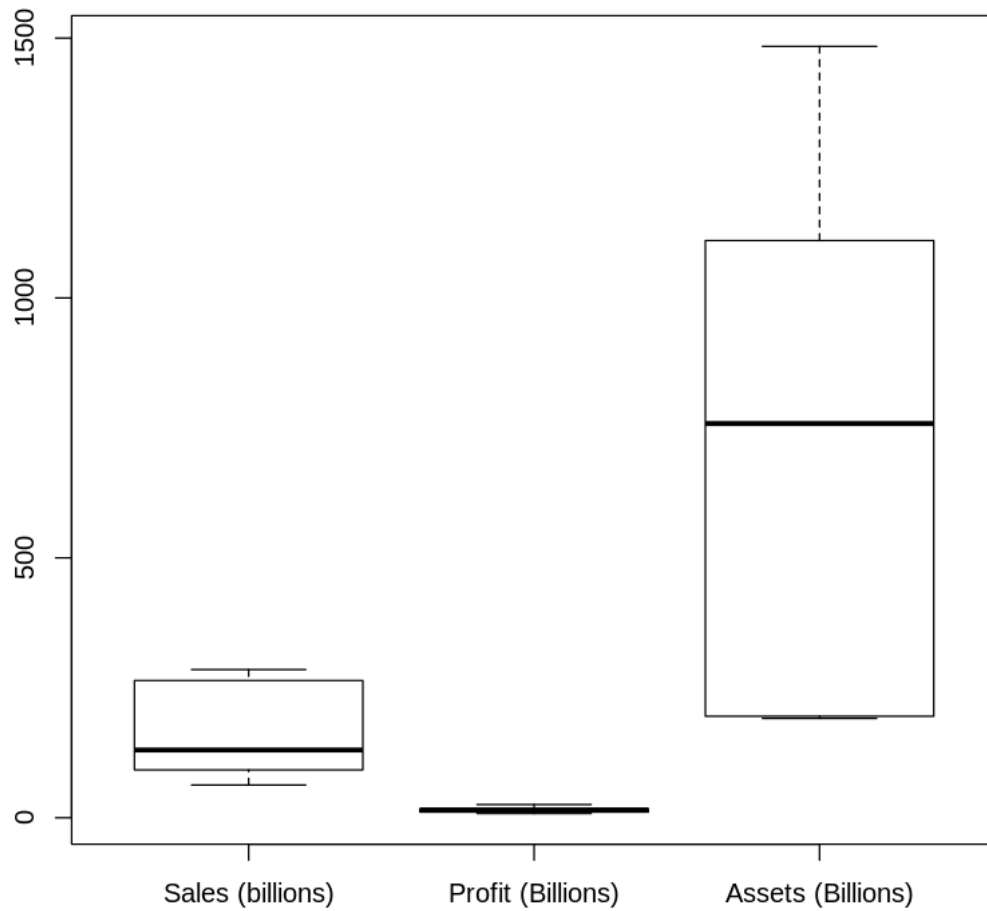
(a) Bivariate scatter plot

```
[2]: plot(data[,2:4])
```
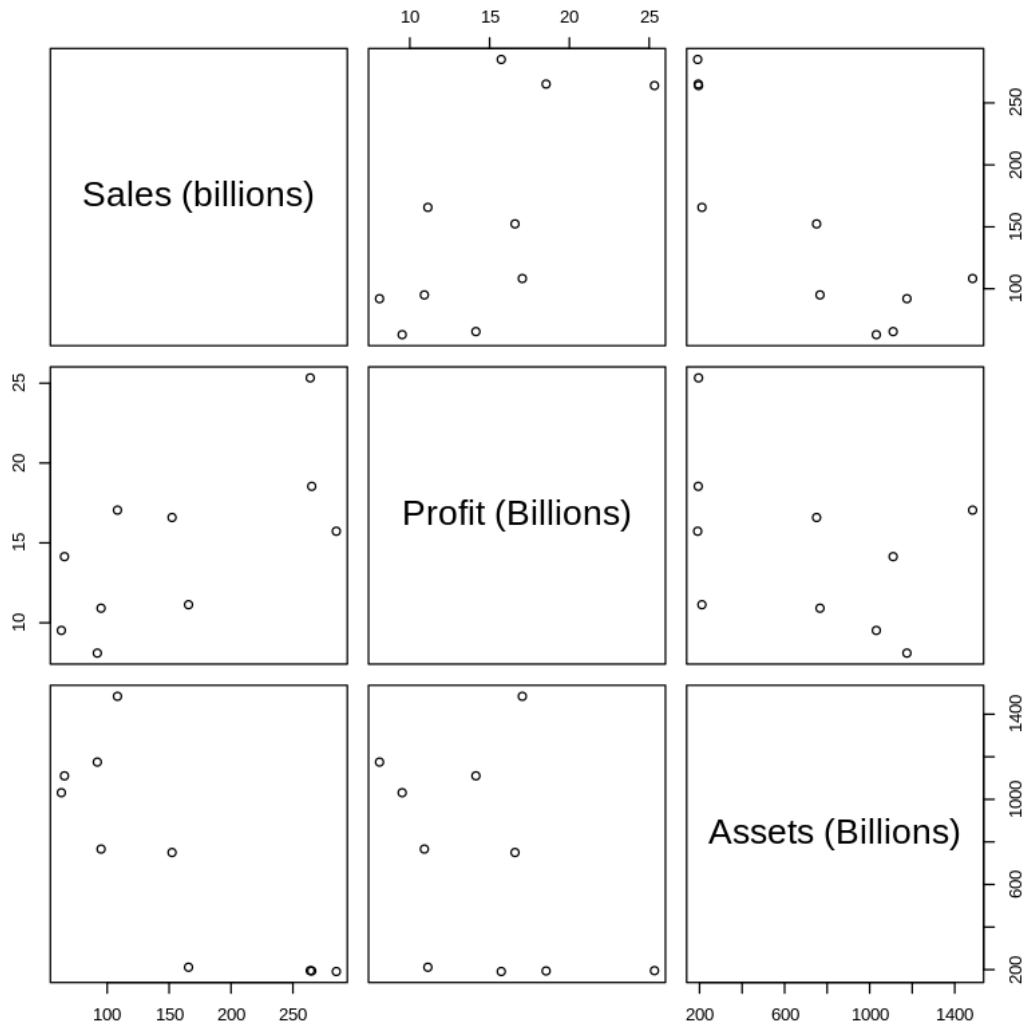
(b) Boxplot

`boxplot(data[,2:4])`

(c)

Scale could be an issue in computing covariance. As we can see in the boxplot above, the values for Assets is much greater and than Sales and Profits. This means when we compute the covariance, the covariance associated with Assets variable would appear to have a much greater magnitude.

(d)

```
[4]: #data matrix
     X <- matrix(c(data$`Sales (billions)`,data$`Profit (Billions)`,data$`Assets␣
      ↪(Billions)`),nrow=10,ncol=3,byrow=F)

     #sample mean vector
     x_bar <- colMeans(X)
     x_bar
```

1. 155.603 2. 14.704 3. 710.911

The sample mean vector is:

$$\vec{x} = \begin{pmatrix} 155.603 \\ 14.704 \\ 710.911 \end{pmatrix} \tag{1}$$

```
[5]: R <- cor(X)
     R
```

	1.0000000	0.6861360	-0.8450549
A matrix: 3 × 3 of type dbl	0.6861360	1.0000000	-0.4229366
	-0.8450549	-0.4229366	1.0000000

The sample correlation matrix is:

$$R = \begin{bmatrix} 1.0000000 & 0.6861360 & -0.8450549 \\ 0.6861360 & 1.0000000 & -0.4229366 \\ -0.8450549 & -0.4229366 & 1.0000000 \end{bmatrix} \tag{2}$$