Assignment 8

Abhishek Murthy

2024-09-22

```r
library(reshape2)
library(corrplot)
```

```
## corrplot 0.92 loaded
```

```r
data <- tips
head(data)
```

```
##   total_bill  tip    sex smoker day   time size
## 1      16.99 1.01 Female     No Sun Dinner    2
## 2      10.34 1.66   Male     No Sun Dinner    3
## 3      21.01 3.50   Male     No Sun Dinner    3
## 4      23.68 3.31   Male     No Sun Dinner    2
## 5      24.59 3.61 Female     No Sun Dinner    4
## 6      25.29 4.71   Male     No Sun Dinner    4
```

```r
str(data)
```

```
## 'data.frame':    244 obs. of  7 variables:
##  $ total_bill: num  17 10.3 21 23.7 24.6 ...
##  $ tip       : num  1.01 1.66 3.5 3.31 3.61 4.71 2 3.12 1.96 3.23 ...
##  $ sex       : Factor w/ 2 levels "Female","Male": 1 2 2 2 1 2 2 2 2 2 ...
##  $ smoker    : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 1 1 1 ...
##  $ day       : Factor w/ 4 levels "Fri","Sat","Sun",..: 3 3 3 3 3 3 3 3 3
## 3 ...
##  $ time      : Factor w/ 2 levels "Dinner","Lunch": 1 1 1 1 1 1 1 1 1 1 1
## ...
##  $ size      : int  2 3 3 2 4 4 2 4 2 2 ...
```

```r
summary(data)
```

```
##    total_bill         tip             sex       smoker      day
## time
##  Min.   : 3.07   Min.   : 1.000   Female: 87   No :151   Fri :19
## Dinner:176
##  1st Qu.:13.35   1st Qu.: 2.000   Male  :157   Yes: 93   Sat :87    Lunch :
## 68
##  Median :17.80   Median : 2.900                          Sun :76
##  Mean   :19.79   Mean   : 2.998                          Thur:62
##  3rd Qu.:24.13   3rd Qu.: 3.562
##  Max.   :50.81   Max.   :10.000
##       size
```

```
##  Min.   :1.00
##  1st Qu.:2.00
##  Median :2.00
##  Mean   :2.57
##  3rd Qu.:3.00
##  Max.   :6.00
```
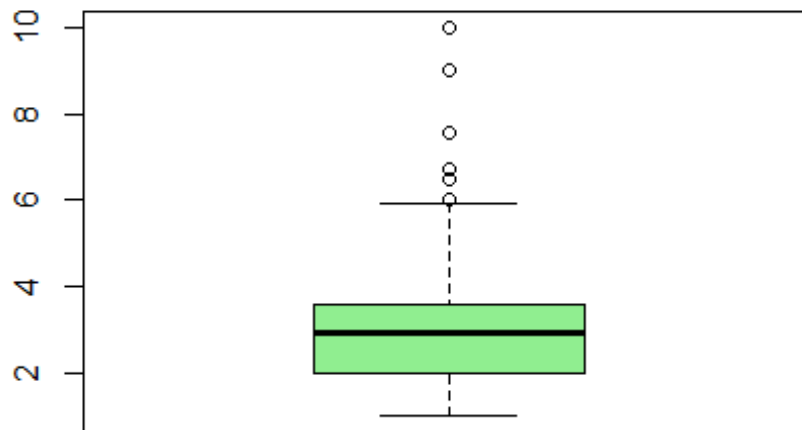
```r
sum(is.na(data))
```

```
## [1] 0
```

```r
hist(data$total_bill, main="Distribution of Total Bill", xlab="Total Bill",
col="lightblue", border="black")
```
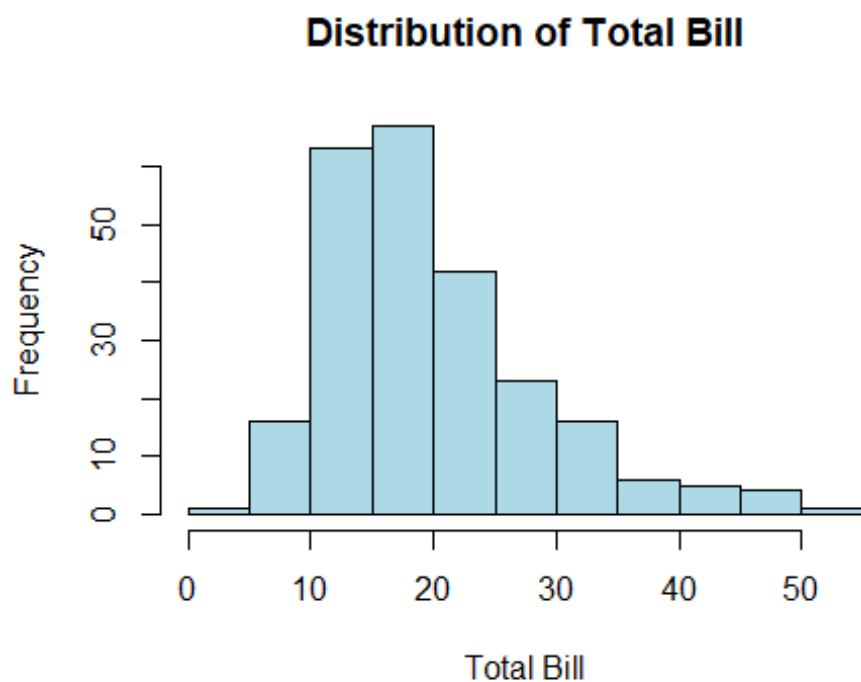


```r
boxplot(data$tip, main="Boxplot of Tip Amount", col="lightgreen")
```
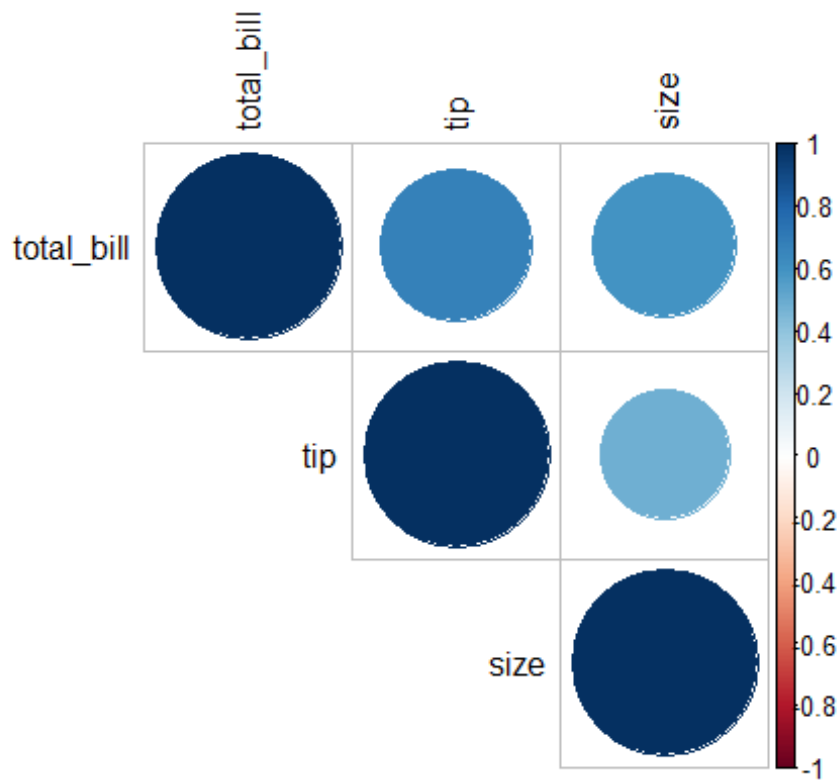
Boxplot of Tip Amount

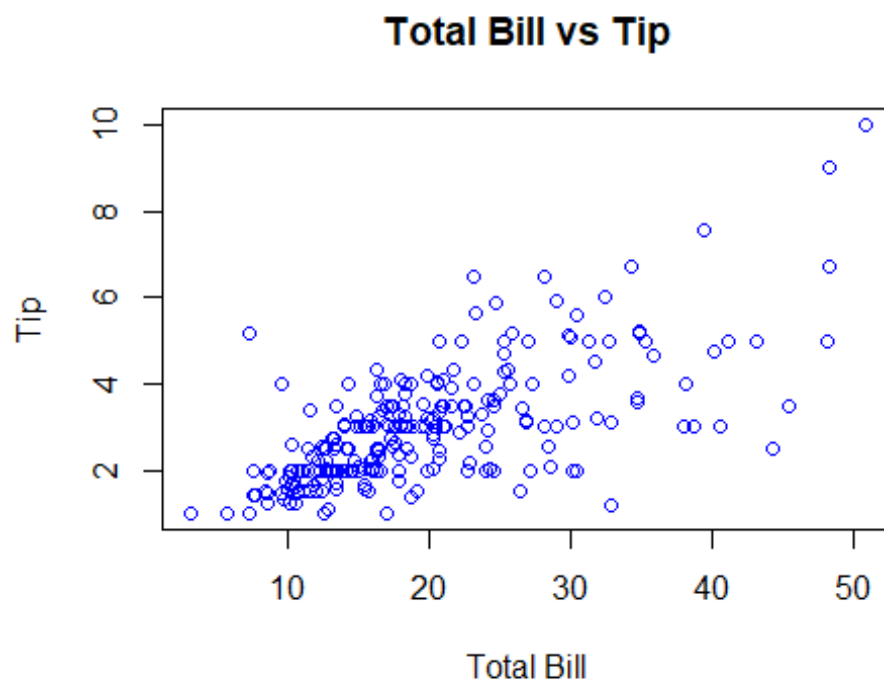```
cor_matrix <- cor(data[, sapply(data, is.numeric)])
print(cor_matrix)

total_bill   tip   size
total_bill   1.0000000 0.6757341 0.5983151
tip   0.6757341 1.0000000 0.4892988
size   0.5983151 0.4892988 1.0000000

corrplot(cor_matrix, method="circle", type="upper", tl.col="black")
```
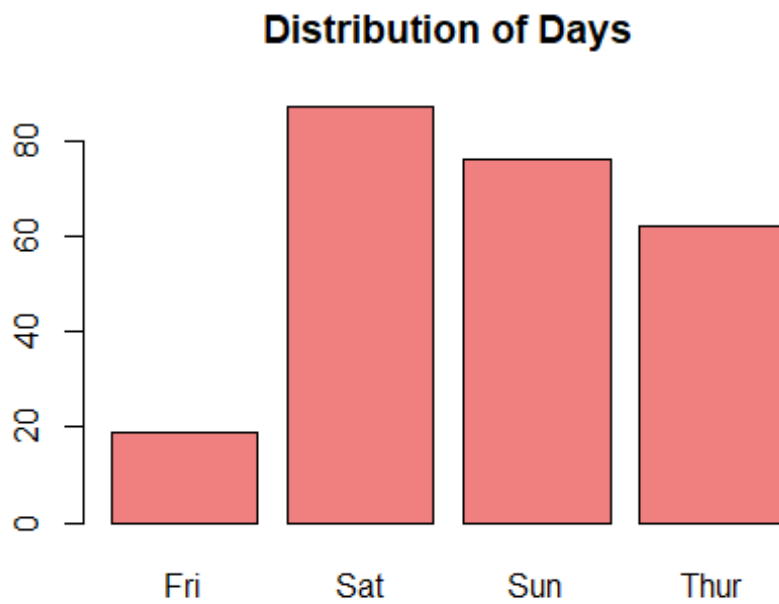
```
plot(data$total_bill, data$tip, main="Total Bill vs Tip", xlab="Total Bill",
ylab="Tip", col="blue")
```

```r
barplot(table(data$day), main="Distribution of Days", col="lightcoral")
```

Distribution of Days


```r
boxplot(tip ~ day, data=data, main="Tip Amount by Day", col="orange")
```

Tip Amount by Day