

Exercise Report

EXERCISE 1

```
# Load the dataset
titanic <- read.csv("./datasets/titanic.csv", header=TRUE, sep=',')
```

```
# Remove the row index column
titanic <- subset(titanic, select=-X)
```

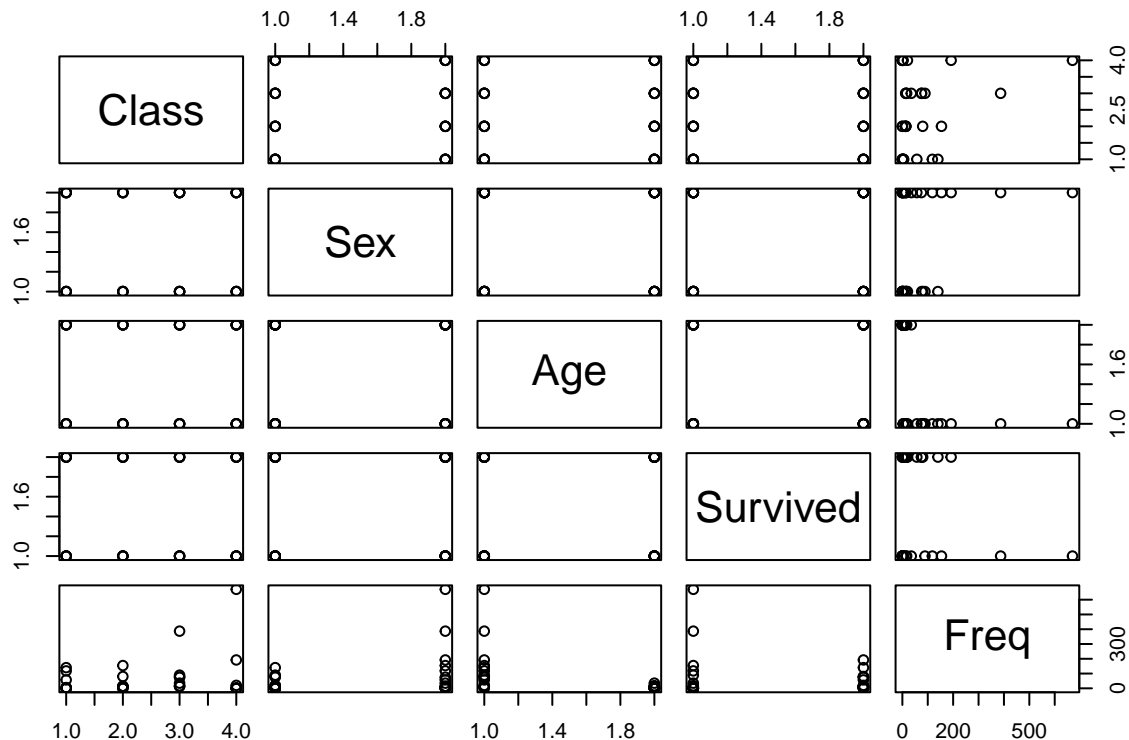
```
# Display the first few rows of the dataset
head(titanic)
```

```
##   Class    Sex  Age Survived Freq
## 1   1st  Male Child      No    0
## 2   2nd  Male Child      No    0
## 3   3rd  Male Child      No   35
## 4  Crew  Male Child      No    0
## 5   1st Female Child      No    0
## 6   2nd Female Child      No    0
```

```
# Display the summary statistics of the dataset
summary(titanic)
```

```
##      Class           Sex           Age           Survived
## Length:32      Length:32      Length:32      Length:32
## Class :character Class :character Class :character Class :character
## Mode  :character Mode  :character Mode  :character Mode  :character
##
##
##
##      Freq
## Min.   : 0.00
## 1st Qu.: 0.75
## Median :13.50
## Mean   :68.78
## 3rd Qu.:77.00
## Max.   :670.00
```

```
# Create a scatterplot matrix of the dataset
plot(titanic)
```



```
dev.off()
```

```
## pdf
## 3
```

```
# Using str() function we can see the structure of the dataset
str(titanic)
```

```
## 'data.frame': 32 obs. of 5 variables:
## $ Class : chr "1st" "2nd" "3rd" "Crew" ...
## $ Sex : chr "Male" "Male" "Male" "Male" ...
## $ Age : chr "Child" "Child" "Child" "Child" ...
## $ Survived: chr "No" "No" "No" "No" ...
## $ Freq : int 0 0 35 0 0 0 17 0 118 154 ...
```

Quantitative (Numerical) Variables:

- Freq: Represents the frequency (number of people)

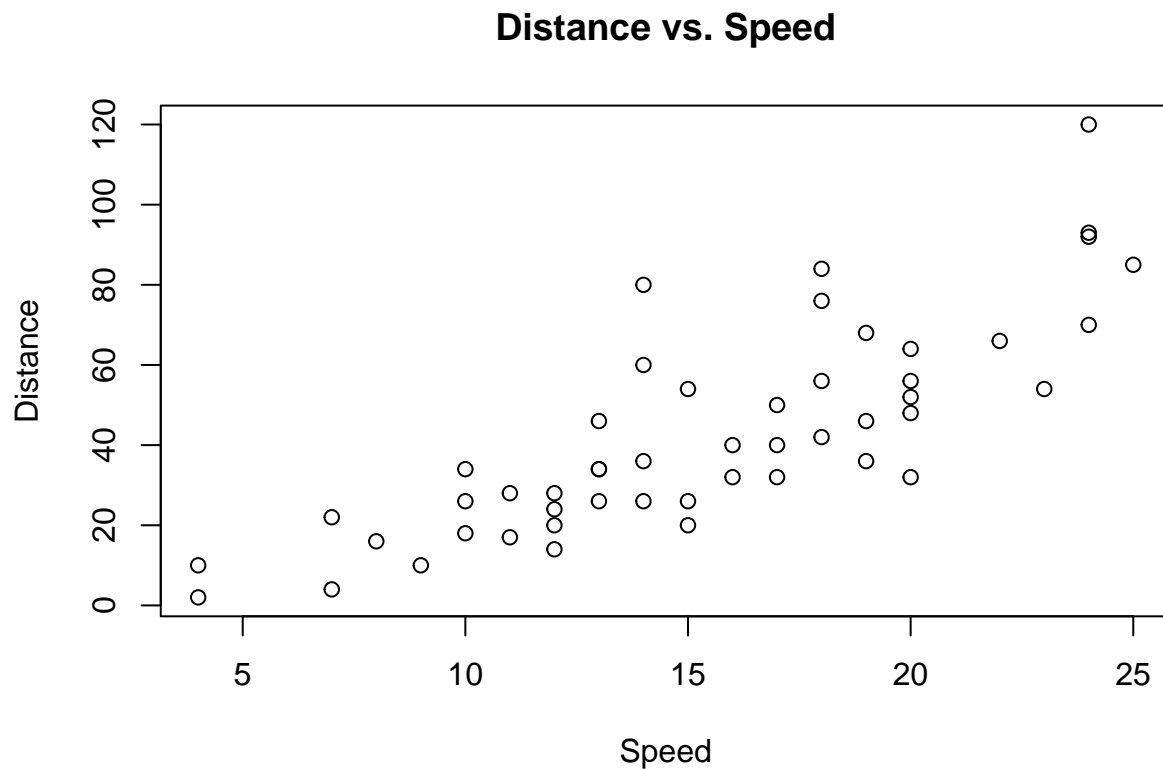
Categorical Variables:

- X: Row index (treated as categorical, despite being an integer)
- Class: Passenger class (e.g., "1st", "2nd", "3rd", "Crew")
- Sex: Gender (e.g., "Male", "Female")
- Age: Age group (e.g., "Child", "Adult")
- Survived: Survival status (e.g., "Yes", "No")

EXERCISE 2

```
# Load the dataset
cars <- read.csv("./datasets/cars.csv", header=TRUE, sep=',')
dev.off()
```

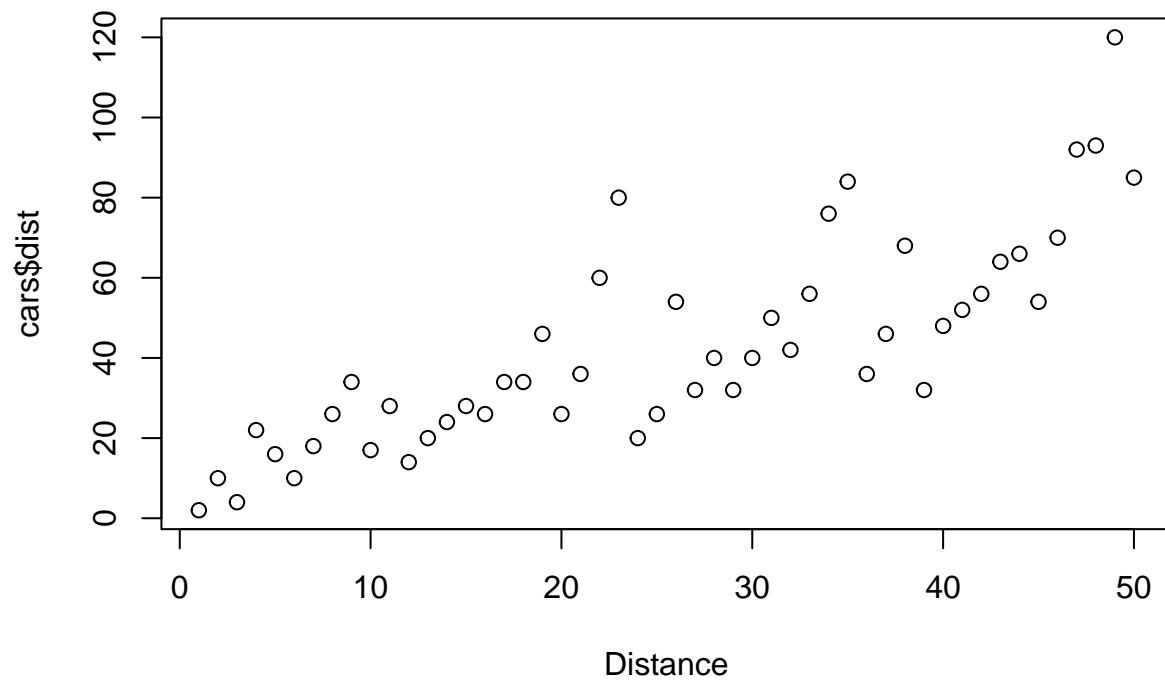
```
## pdf
## 3
# Make a plot of the distance field in terms of the speed field (use the $ syntax)
plot(cars$speed, cars$dist, xlab="Speed", ylab="Distance", main="Distance vs. Speed")
```



```
dev.off()

## pdf
## 3
# Create a histogram of the distance field
plot(cars$dist, xlab="Distance", main="Histogram of Distance")
```

Histogram of Distance



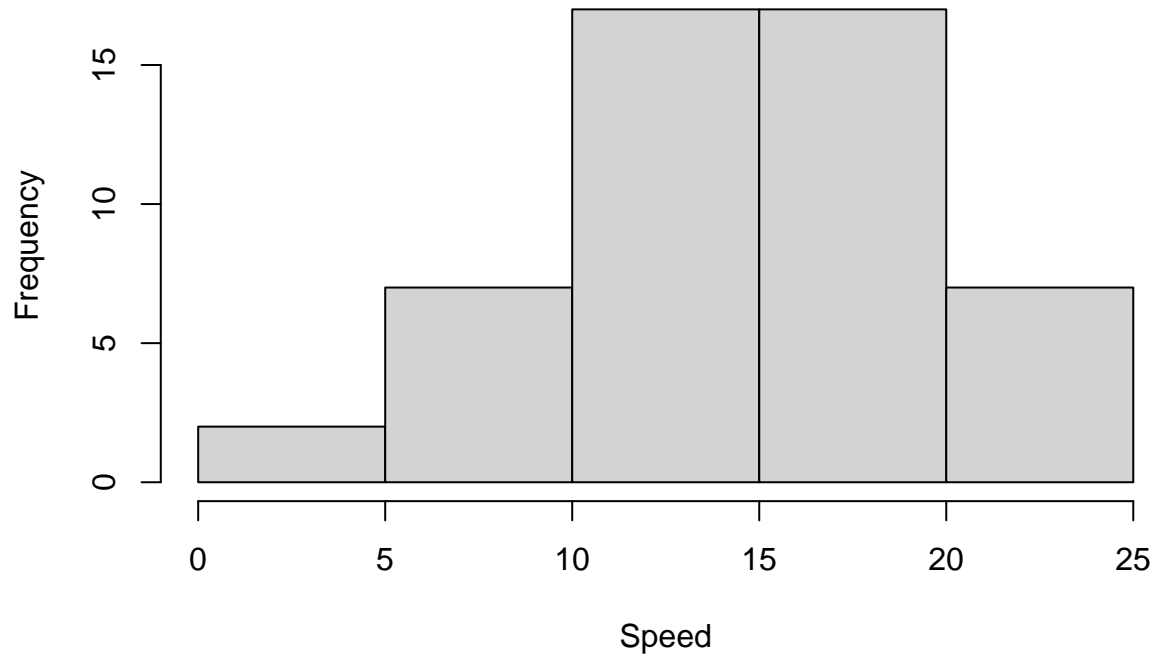
```
dev.off()
```

```
## pdf  
## 3
```

```
# Create a histogram of the speed field
```

```
hist(cars$speed, xlab="Speed", main="Histogram of Speed")
```

Histogram of Speed



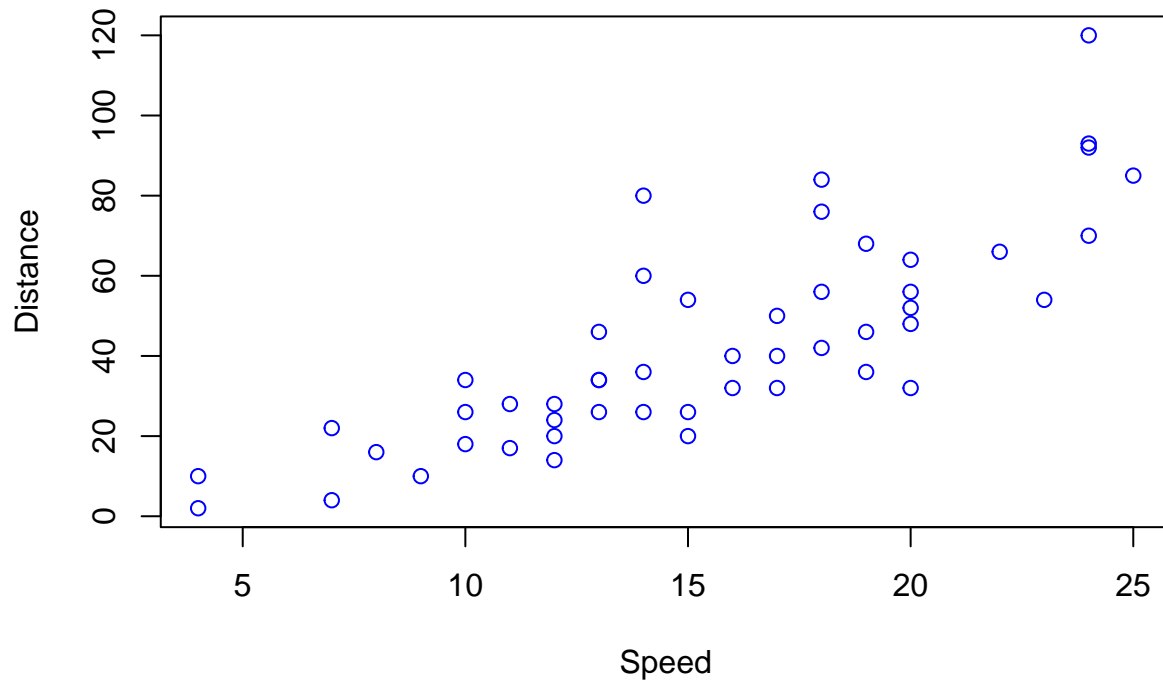
```
dev.off()
```

```
## pdf  
## 3
```

```
# Save the scatterplot of the distance field in terms of the speed field to a PDF file  
pdf("./documents/exercice2/ex2_distance_vs_speed_plot.pdf")
```

```
# Scatterplot of Distance vs Speed with modified title, axis labels, and color  
plot(cars$speed, cars$dist,  
      xlab="Speed",  
      ylab="Distance",  
      main="Distance vs Speed",  
      col="blue")
```

Distance vs Speed



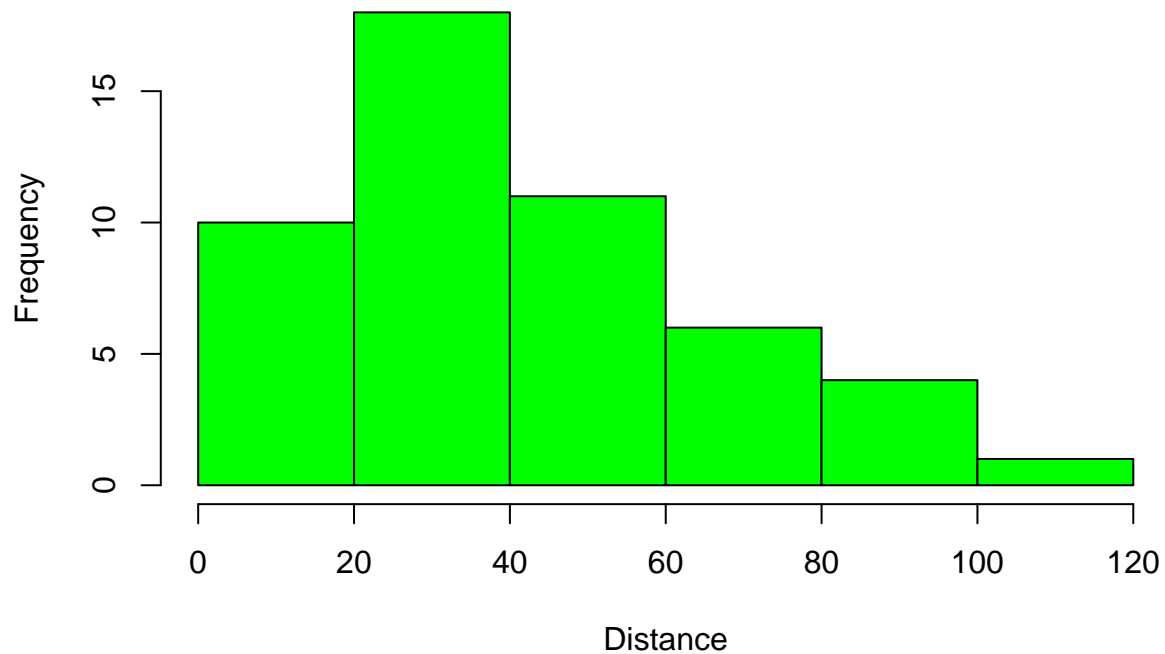
```
dev.off()
```

```
## pdf  
## 3
```

```
# Save the histogram of the distance field to a PDF file  
pdf("../documents/exercice2/ex2_histogram_distance.pdf")
```

```
# Histogram of Distance with modified title, axis labels, and color  
hist(cars$dist,  
      xlab="Distance",  
      main="Histogram of Distance",  
      col="green")
```

Histogram of Distance



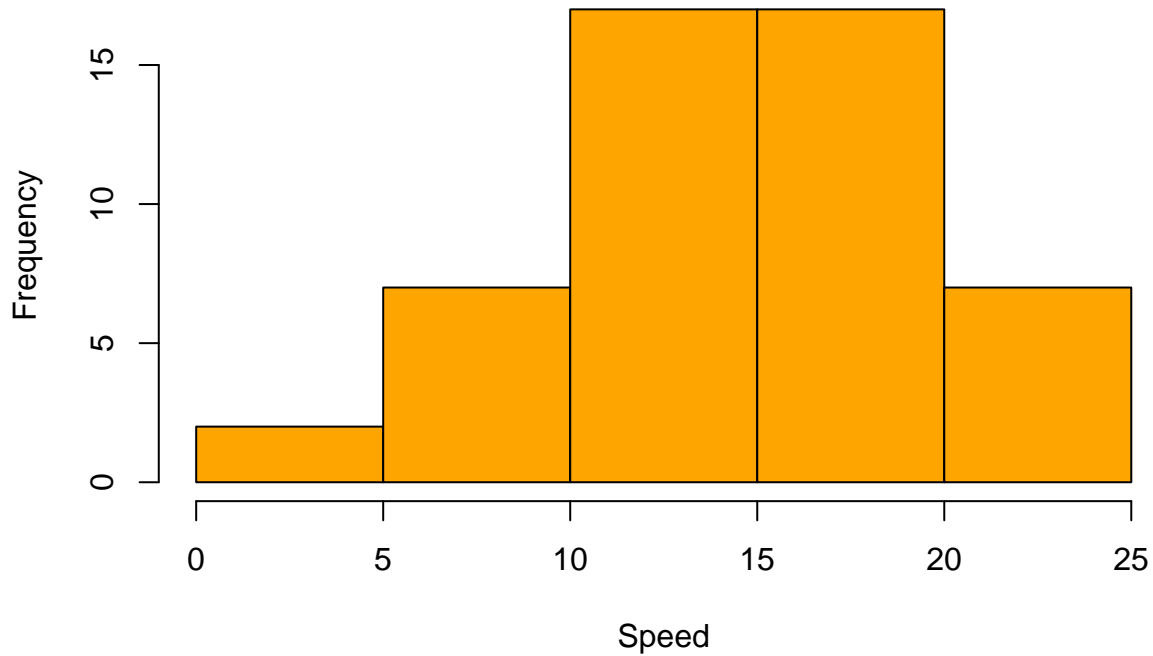
```
dev.off()
```

```
## pdf  
## 3
```

```
# Save the histogram of the speed field to a PDF file  
pdf("./documents/exercice2/ex2_histogram_speed.pdf")
```

```
# Histogram of Speed with modified title, axis labels, and color  
hist(cars$speed,  
      xlab="Speed",  
      main="Histogram of Speed",  
      col="orange")
```

Histogram of Speed



```
dev.off()
```

```
## pdf  
## 3
```

EXERCISE 3

```
# Load the dataset  
cars <- read.csv("./datasets/cars.csv", header=TRUE, sep=',')
```

```
# Remove the first column of the cars data frame  
cars <- cars[, -1]
```

```
# Construct a new data frame  
new_cars <- data.frame(speed = c(21, 34), dist = c(47, 87))
```

```
# Add the constructed data frame to the cars data frame  
cars <- rbind(cars, new_cars)
```

```
# Sort the data in the resulting dataset by column speed (ascending)  
cars <- cars[order(cars$speed), ]
```

```
# Write the resulting dataset to a CSV file  
write.csv(cars, file = "./datasets/cars_sorted.csv", row.names = FALSE)
```