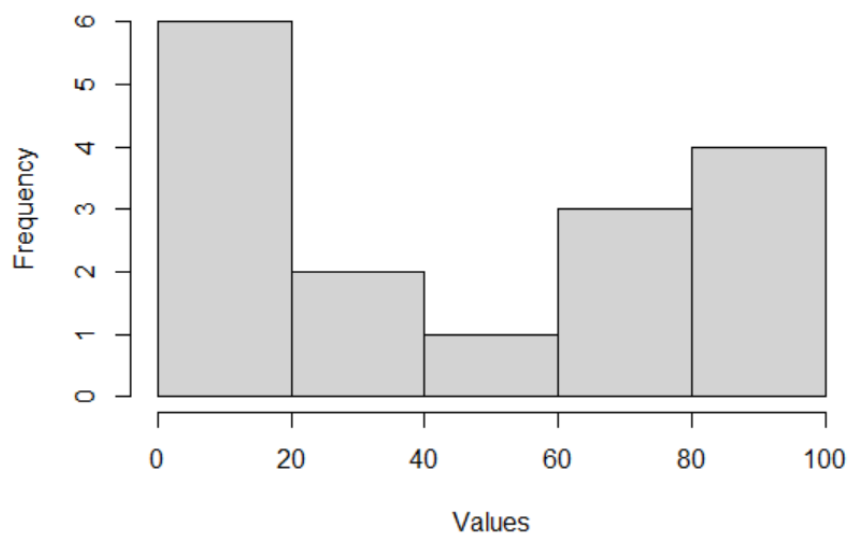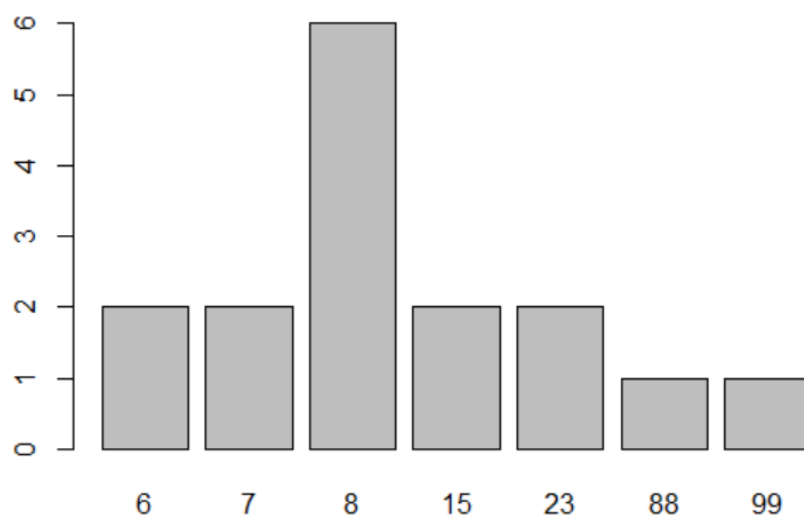## Distribution of Numeric Column

Frequency of Categories


```
 1  install.packages("readr")
 2  library(readr)
 3  getwd()
 4  dataset <- read.csv('mall.csv')
 5  str(dataset)
 6  head(dataset)
 7  summary(dataset)
 8  subset_data <- dataset[dataset$Column_Name > 50, ]
 9  dataset$new_variable <- dataset$old_variable * 2
10  filtered_data <- subset(dataset, Column_Name == "Value")
11  mean_value <- mean(dataset$Numeric_Column)
12  median_value <- median(dataset$Numeric_Column)
13  sd_value <- sd(dataset$Numeric_Column)
14  table(dataset$Categorical_Column)
15  hist(dataset$Numeric_Column, main = "Distribution of Numeric Column", xlab = "Values")
16  barplot(table(dataset$Categorical_Column), main = "Frequency of Categories")
```

Regression Analysis
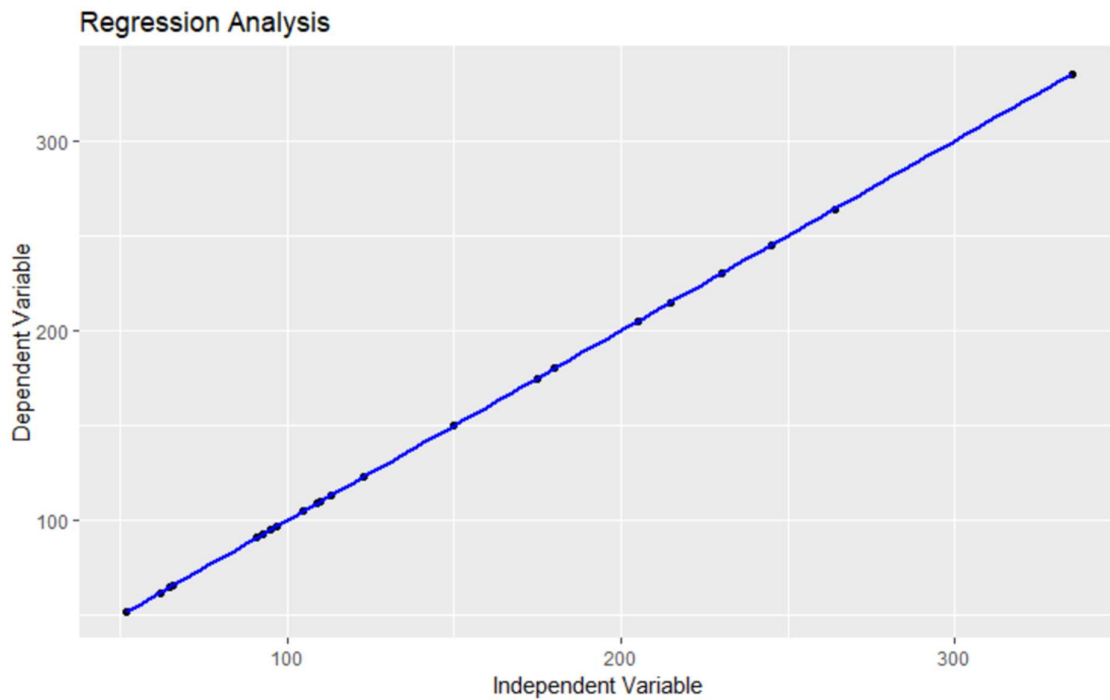

```r
1   library(ggplot2)
2   data <- read.csv("mtcars.csv")
3   # Perform linear regression
4   linear_model <- lm(hp ~ mpg, data = data)
5   # Perform multiple regression
6   multiple_model <- lm(hp ~ mpg + disp, data = data)
7   # Perform polynomial regression
8   poly_model <- lm(hp ~ poly(mpg, degree = 2), data = data)
9   # Evaluate model performance
10  summary(linear_model)
11  summary(multiple_model)
12  summary(poly_model)
13  # Visualize regression results
14  ggplot(data, aes(x = hp, y = hp)) +
15    geom_point() +
16    geom_smooth(method = "lm", se = FALSE, color = "blue") +
17    labs(title = "Regression Analysis", x = "Independent Variable", y = "Dependent Variable")
18
```