STA 3100 Programming With Data in R: Data Visualization

Data Visualization in R

Introduction

Data visualization is the process of creating visual representations of data. It is a powerful tool for understanding and communicating data. Data visualization can help to identify patterns, trends, and correlations in data that would be difficult to detect in other forms. In this lecture, we will discuss the basics of data visualization in R.

Basic Plotting

The simplest way to create a plot in R is with the `plot()` function. This function takes two arguments: a vector of x values and a vector of y values. For example, the following code creates a simple scatter plot of two vectors:

```R
# Start of Code
x <- c(1,2,3,4,5)
y <- c(2,4,6,8,10)
plot(x,y)
# End of Code
```

The `plot()` function also has several optional arguments that can be used to customize the plot. For example, the `type` argument can be used to specify the type of plot to create (e.g. `type = "l"` for a line plot).

Labeling Plots

It is important to label plots so that they are easy to interpret. The `xlab` and `ylab` arguments can be used to add labels to the x-axis and y-axis, respectively. For example, the following code adds labels to the previous plot:

```R
# Start of Code
x <- c(1,2,3,4,5)
y <- c(2,4,6,8,10)
plot(x,y, xlab = "X Values", ylab = "Y Values")
# End of Code
```

The `main` argument can be used to add a title to the plot.

Multiple Plots

The `par()` function can be used to create multiple plots in a single figure. The `mfrow` argument can be used to specify the number of rows and columns in the figure. For example, the following code creates a 2x2 figure with four plots:

```R
# Start of Code
par(mfrow = c(2,2))
plot(x,y)
plot(x,y, type = "l")
plot(x,y, xlab = "X Values", ylab = "Y Values")
plot(x,y, main = "My Plot")
# End of Code
```

Practice Multiple Choice Questions

Q1. What is data visualization?

A. The process of creating visual representations of data

B. The process of analyzing data

C. The process of collecting data

D. The process of organizing data

Answer: A. The process of creating visual representations of data