

STA 3180 Statistical Modelling: Data Visualization

Data Visualization in STA 3180 Statistical Modelling

Data visualization is a powerful tool for exploring data, understanding relationships between variables, and communicating results. It is a key component of statistical modelling and can help to provide insights that would otherwise be difficult to uncover. In this lecture, we will discuss the basics of data visualization, including how to create visualizations using R, and how to interpret them.

What is Data Visualization?

Data visualization is the process of creating graphical representations of data. It is used to explore and understand data, identify patterns and trends, and communicate results. Data visualizations can be used to show relationships between variables, compare distributions, and illustrate trends over time.

Types of Data Visualizations

There are many different types of data visualizations, including line graphs, bar charts, scatter plots, histograms, and box plots. Each type of visualization has its own advantages and disadvantages, and should be chosen based on the type of data being visualized and the story it is trying to tell.

Creating Visualizations with R

R is a powerful programming language for data analysis and visualization. It has a wide range of packages and functions that make it easy to create visualizations. To create a visualization in R, you first need to load the necessary packages and data. Then, you can use the `ggplot2` package to create the visualization.

Start of Code

```
# Load packages
library(ggplot2)

# Load data
data <- read.csv("data.csv")

# Create visualization
ggplot(data, aes(x = x_variable, y = y_variable)) +
  geom_point()
```

End of Code

Interpreting Visualizations

When interpreting visualizations, it is important to look for patterns and trends. Look for clusters, outliers, and changes over time. Also, consider the context of the visualization and how it relates to the data.

Practice Multiple Choice Questions

Q1. Which of the following is NOT a type of data visualization?

- A. Pie chart
- B. Box plot
- C. Bubble chart
- D. Word cloud

Answer: D. Word cloud