

PROLOGUE

The first part of the book discusses the importance of understanding the context of the data being analyzed. It emphasizes the need for a clear research question and the selection of appropriate statistical methods. The author also discusses the challenges of working with large datasets and the importance of data cleaning and preprocessing.

Method	Advantages	Disadvantages
Linear Regression	Simple and easy to interpret	Assumes a linear relationship between variables
Logistic Regression	Used for binary classification	Assumes a log-linear relationship between variables
Decision Trees	Can handle non-linear relationships	Can be prone to overfitting
Support Vector Machines	Can handle high-dimensional data	Can be computationally expensive
Neural Networks	Can learn complex, non-linear relationships	Requires large amounts of data and computational power

KINGSTON

