Introduction to Statistical Learning Book Exercises

Kevin Sullivan

2025-01-14

Chapter 2: Statistical Learning

Chapter Topics:

1. Prediction
2. Inference
3. Parametric Methods
4. Non-Parametric Methods
5. Trade-Off Between Prediction Accuracy and Model Interpretability
6. Supervised Vs. Unsupervised Learning
7. Assessing Model Accuracy
8. Measuring the Quality of Fit
9. Bias - Variance Trade-Off
10. Classification
11. K-Nearest Neighbors

Practice Question: Load and Perform Exploratory Data Analysis

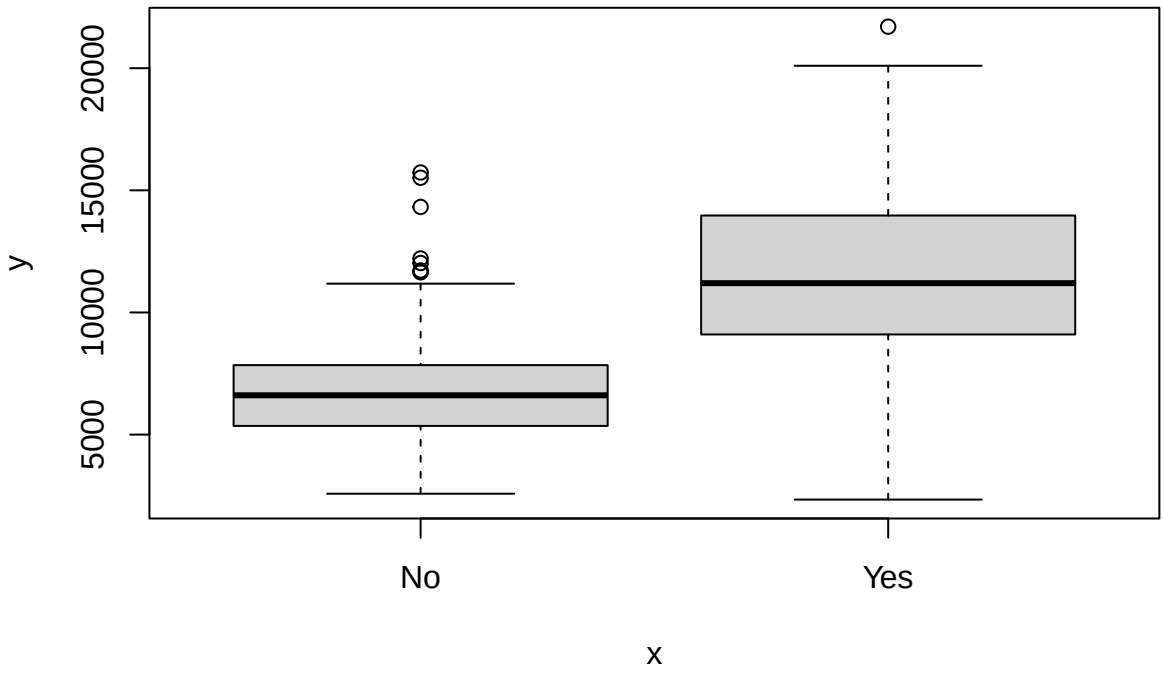
```
# Load Data
chap_two_data = College
```

```
head(chap_two_data)

summary(chap_two_data)
```

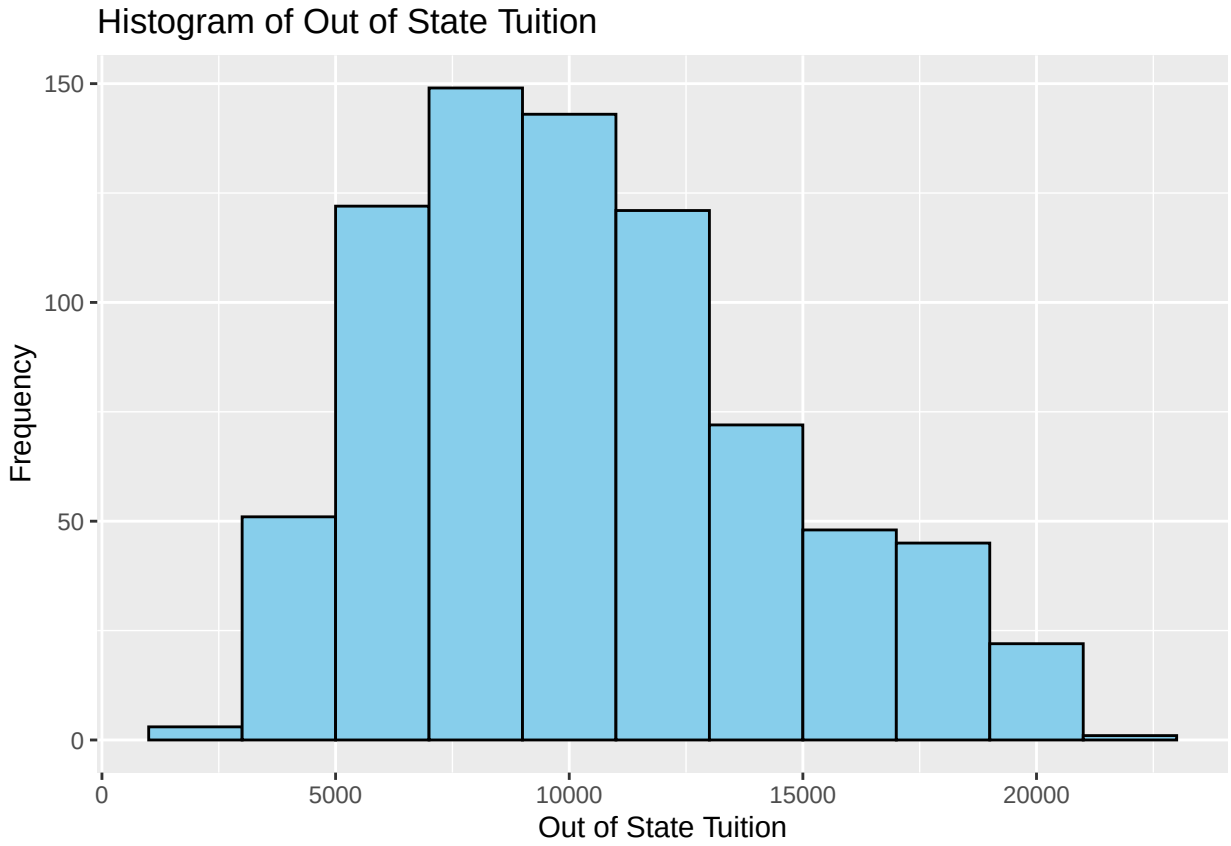
```
# Exploratory Data Analysis

plot(chap_two_data$Private, chap_two_data$Outstate)
```




```
Elite = rep("No", nrow(chap_two_data))
Elite[chap_two_data$Top10perc > 50] = "Yes"
Elite <- as.factor(Elite)
college = data.frame(chap_two_data, Elite)

college %>%
  ggplot(aes(x = Outstate)) +
  geom_histogram(binwidth = 2000, fill = "skyblue", color = "black") +
  labs(title = "Histogram of Out of State Tuition", x = "Out of State Tuition", y = "Frequency")
```



Chapter 3: Linear Regression

Chapter Topics:

1. Simple Linear Regression
2. Assessing Accuracy of Coefficient Estimates
3. Assessing Accuracy of the Model
4. Multiple Linear Regression
5. Qualitative Predictors
6. Potential Problems
7. Comparison of Linear Regression and KNN